A thesis submitted to the University of Gloucestershire
The Business School
in accordance with the requirements of the degree of
Doctor of Business Administration

Quality of Stock Price Predictions in Online Communities –

Groups or Individuals?

Author	Tobias Endress
Student Number	s0914570
E-Mail	tobias@endress.info
Submitted	28/03/2017

Abstract

Group decision-making and equity predictions are topics that are interesting for academic research as well as for business purposes. Numerous studies have been conducted to assess the quality of forecasts by financial analysts, but in general these studies still show little evidence that it is possible to generate accurate predictions that in the long run create, after transaction costs, profits higher than the market average. This thesis investigates an alternative approach to traditional financial analysis. This approach is based on Internet group decision-making and follows the suggestion that a group decision is better than the decision of an individual. The research project follows a mixed-methods approach in the form of a sequential study with a field experiment. Different groups—consisting of lay people, but also financial professionals—were formed purposefully in different group designs to generate equity forecasts. The field experiment was conducted following an e-Delphi approach with online questionnaires, but also in-depth interviews with all participants. Data from financial analysts was used to compare the predictions from the groups with actual results of share prices.

The data from the experiment suggests that there are different variables, in terms of the individual characteristics of the participants, which indicated significant impact on the quality of equity predictions. The predictions of some participants (e.g. “PID-S-plus” rated participants) are apparently of significantly higher accuracy. The findings from the study indicate that intuition plays a significant role in the decision-making process not only for lay people, but also for financial analysts and other financial professionals. However, there are observable differences in the intuitive decision-making of lay people and experts. While it was possible to observe that intuition is interpreted as “random guess” by poor predictors, it was found that good predictors base their intuition on several factors—even including fundamental and macroeconomic considerations. The findings of the experiments led to an explanatory model that is introduced as the ‘Deliberated Intuition’ Model. The model of

deliberated intuition which is proposed here views prediction as a process of practice which will be different for each individual. The model proposes that a predictor will decide, consciously or semi-consciously, when they feel ready to rely on gut-feeling, or to undertake more analysis. Generally, it appears to contribute to a good prediction to think about the problem in different ways and with various techniques. The experiment indicated that (online-) groups are not per se better than individuals. The Deliberated Intuition Model might help to prepare better group settings and improve prediction quality. Apparently a combination of rational and intuitive techniques leads to the best prediction quality.

Author's Declaration

I declare that the work in this thesis was carried out in accordance with the regulations of the University of Gloucestershire and is original except where indicated by specific reference in the text. No part of the thesis has been submitted as part of any other academic award. The thesis has not been presented to any other education institution in the United Kingdom or overseas.

Any views expressed in the thesis are those of the author and in no way represent those of the University.

Signed ………………………………………… Date ……………………………….

Table of Contents

Abbreviations

AUM	Assets under Management
BEst	Bloomberg Estimates
CA	Conversation Analysis
CI	Collective Intelligence
DA	Disclosure Analysis
DCF	Discounted Cash Flow
DF	Degree of Freedom (statistics)
DSS	Decision Support Systems
EMH	Efficient Market Hypothesis
EPS	Earnings per Share
F	Frequency (SPSS-Output)
GDSS	Group Decision Support Systems
IEM	Iowa Electronic Market
IS	Information Systems
NGT	Nominal Group Technique
p-value	Calculated Significance Value (in SPSS-Output sometimes “Signif.”)
PA	Prediction Accuracy (Percentage of Correct Predictions)
PE	Price-Earnings Ratio
PM	Prediction Market
TP	Target Price (Stock Price Prediction)
TV	Television
S&P 500	Standard & Poor's 500 (stock index)

List of Figures

List of Tables

1 Introduction: Stock Price Predictions in Online Communities

Equity research is a topic that is interesting for academic research as well as for business purposes. The work of academics who focus on financial markets and of business financial analysts is of 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). Nevertheless, there are many doubts about the quality of their work. A myriad of studies have already been conducted to assess the quality of the resulting forecasts by financial analysts (Aiolfi, Rodriguez, & Timmermann, 2009; Bolliger, 2004; Clement, 1999; Fleischer, 2005; Stanzel, 2007), but in general these studies still show little evidence that it is possible to generate accurate predictions that in the long run create, after transaction costs, profits higher than the market average (Malkiel, 2007; Stanzel, 2007). Therefore, it seems necessary to conduct further analysis and develop more reliable assessment approaches to identify good investment ideas as early as possible.

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 a group decision approach (Kaplan, 2001), which is used by several special interest (stock trading) communities on the Internet. This approach follows the idea that a group decision is better than the decision of an individual (Page, 2008b; Sunstein, 2008; Surowiecki, 2005).

One should also note that some authors doubt that groups can decide better than an expert (e.g., Dueck, 2015; F. B. Simon, 2013); 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). Others, such as the philosopher Friedrich Nietzsche (1989), wrote that madness is rare in individuals, but he regarded it as the rule in groups, while Gustave Le Bon regarded crowds as “organisms”, but argued that they can never attain a high degree of intelligence (2009). Literature describes a wide range of issues in the context of group-decision-making, e.g. the anchoring effect (Tversky & Kahneman, 1974), conformity, group pressure (Asch, 1956; Milgram, 1964), or higher risk propensity (Nijstad, 2009; Stoner, 1961). These two contrasting, but equally compelling views—regarding groups as “smarter” or groups as unintelligent—rest on how the respective author views the “operation” of the group, and lead to an examination of the issues that influence such operations.

Kaplan (2001) described a prototype of a system for forecasting stock prices using collective intelligence with quite positive and promising experience with a test run of his system. Kaplan's paper still leaves issues open for further research. He suggests “conducting a test on a much larger scale, and experimenting with variations of the [collective intelligence] CI processing algorithms to identify those that are most effective” (2001, p. 6). Since his suggestion of further tests on a larger scale, the approach has been used in practice by several investment communities, so that observation and examination of these portals (e. g., marketocracy.com, predictwallstreet.com, or sharewise.com) could lead to answering further questions such as: under what circumstances, including the mechanisms driving the decision-making process, would a remote group like an Internet community outperform the equity research forecast accuracy of an individual financial analyst? Kaplan (2001) did not fully disclose in his paper the algorithm and process of the group decision-making methods used. Accordingly, a first step towards creating a better understanding is to conduct a trial with a clearly defined process in a more controlled environment.

This thesis is structured in the following main sections: “Literature Review”, “Research Methodology and Methods”, “Pilot Experiment”, “Main Experiment”, “Contribution to Knowledge and Business Practice” and “Synopsis and Conclusion”. The literature review section provides background information on traditional equity research methods and presents the existing evidence in the context of group decision-making. The literature suggests that there is still limited knowledge about the underlying qualitative factors and even less knowledge about the accuracy and quality of financial forecasts by online groups. However, there is a body of evidence about factors that contribute and foster good group decisions. Apparently, numerous existing online communities, in particular communities which focus on equity predictions, fail to apply this existing knowledge in an appropriate way. This research project aims to link previous knowledge, new insights and practical application. Overall aim of the study: to explore, analyse and compare the quality of equity predictions of individuals and groups who are using the Internet in order to build theory of process. The research methodology and methods section introduces the questions and objectives, as well as the selection of applied research tools and methods. Derived from existing knowledge the following research questions emerged:

- Research question 1:

  How do the recommendations of an Internet group perform in comparison to the recommendations of an individual expert (financial analyst)?

- Research question 2:

  How does the feedback loop of an e-Delphi process affect the prediction quality of an Internet group making equity predictions?

- Research question 3:

  What are the underlying key mechanisms, of the individual and of the group, that influence the decision-making process, and how might the decision-making process in existing online communities be improved?

The research project designed to address these questions follows a mixed-methods approach in the form of a sequential study. While a purely positivist approach would have been appropriate to address questions 1 and 2, a constructivist perspective was found helpful in gaining a more holistic understanding of the underlying mechanisms.

The mixed-methods approach is based on an experimental research design. In order to validate and improve the research design a pilot test on the operation of the online process for the proposed research was conducted. The pilot experiment demonstrated the feasibility of the research project and its ability to address the research questions. Furthermore, the pilot experiment also provided an indication of how the research design might be improved. The key elements learned from the pilot experiments can be categorized as follows:

- Adjust group design and feedback loop
- Assessment of the participants
- Enhancements of the online questionnaire

The pilot run of the experiment provided a few indications that it might be possible to facilitate a process with an online group that made it possible to make–in certain situations and with careful group design–predictions that are superior to predictions by experts.

The main experimental section presents the approach and findings of the sequential mixed-methods study following the pilot run. Quantitative data analysis was conducted in a sequential approach: a univariate analysis and secondly a multi-criteria analysis. Building on the findings from the pilot run, the main experiment was conducted using a bigger sample, a longer period, more shares and a wider range of different group designs. An extensive discussion of the hypotheses developed is included in this section.

As expected, there is a certain degree of random walk in the process of predicting stock prices. Nevertheless, there are factors that appear to improve predictive accuracy.

Some of these factors are inherent in the personality of the predictor. E.g. their score on the PID-scale or gender. Some other factors can have influences, some—like education level—more long term; and some may be facilitated directly with the decision-making process.

The findings from the study indicate that intuition plays a significant role in the decision-making process not only for lay people, but also for financial analysts and other financial professionals. Still, there are observable differences in the intuitive decision-making of lay people and experts. While it was possible to observe that intuition is interpreted as “random guess” by poor predictors, it could be seen that good predictors base their intuition on several factors—even including fundamental and macroeconomic considerations. The findings of the experiments led to an explanatory model that is introduced as the ‘Deliberated Intuition’ Model. Generally, it appears to contribute to a good prediction to think about the problem in different ways and with various techniques. Apparently a combination of rational and intuitive techniques leads to the best predictive quality.

2 Literature Review

This literature review is primarily intended to identify qualitative and quantitative research evidence and aspects of group decision-making, particularly with regard to Internet communities which focus on stock trading issues as their basis.

The literature review is conducted in four parts, applying a combined review approach. This combined approach allows us to introduce and discuss separately the context and the basic ideas and then the interdependence of these key factors. The first part introduces the traditional ideas and theoretical background of traditional equity research approaches, as well as the general constraints of traditional equity research methods in order to clarify the background and context of the study. The next two sections, “Group Decision-Making” and “Decision Support Systems”, are reviewed using a critical narrative review approach (Baumeister & Leary, 1997); this allows an overview of the work as well as identifying key theories, concepts, and ideas, and highlighting critical issues regarding group decision-making. These ideas and concepts are used to identify criteria that might influence decision-making in groups. The identified criteria form a basis for assessing the quality of decisions made by Internet groups about the development of stock prices and the respective “Buy” and “Sell” recommendations of these groups. The fourth part, “Internet Group Decision-Making”, was conducted using an approach that systematically screens literature databases to show their relevance, giving an overview of the existing knowledge in the field, and identifying gaps in the literature (Randolph, 2009). Another aim of this part is to introduce the concept of generating investment ideas and stock price predictions in Internet groups, and to suggest the existing body of knowledge in this field as a starting point for further research. As Tetlock points out: “We know that in so much people want to predict—politics, economics, finance, business, technology, daily life—predictability exists, to some degree, in some circumstances. But there is so much else we don't know” (Tetlock & Gardner, 2015, p. 16). And it might be easy to agree with him that “[f]or scientists, not knowing is exciting. It's an opportunity to discover; the more that is unknown, the greater the opportunity. Thanks to frankly quite amazing lack of rigor in so many forecasting domains, this opportunity is huge” (Tetlock & Gardner, 2015, p. 16). Indeed, it seems that this observation applies also, to some extent, to financial forecasts. While there are at least some quantitative follow-up mechanisms for professional analysts in place (e.g. StarMine Monitor), there is still very little knowledge about the underlying qualitative factors and even less knowledge about the accuracy and quality of forecasts by online groups (with lay people's predictions). This research project contributes to the body of knowledge in this context, in particular by gaining a better understanding of underlying mechanisms and influential factors in the context of equity predictions by financial analysts, lay people, and online groups.

2.1 Evidence in Literature

Even though some recent authors argue that they have observed a “new collaborative economy” (Chase, 2015) the literature suggests that the idea of letting a group decide is not really as new as the popularity of the book *The Wisdom of Crowds* (Surowiecki, 2005) might indicate. In fact, there has been a lot of research about group versus individual decision-making. This “new” approach might contain or combine parts of established group individual decision-making-procedures such as the Delphi methodology, the nominal group technique, prediction markets, Internet decision-making, the social psychology of groups, and group support systems, which are quite well covered by academic research. That is why in the following sections the related ideas and theories about decision-making are discussed. The relevant research conducted in these fields is examined for attributes that could possibly affect the decision-making process of special interest communities on the Internet. Another reason is to introduce basic concepts and theories to build on this foundation later on in the study. But before continuing to discuss group decision-making further, the traditional approach to equity research, the context of this study, has to be introduced so as to allow benchmarking the one against the other.

2.2 Equity Research

Nils Bohr once joked that “prediction is very difficult, especially about the future” (as cited in Ellis, 1977, p. 431). In fact, such an observation applies specifically to the process of forecasting stock prices. Nevertheless, thousands of people all over the world rely on predictions every day when they consider investment decisions. It is a major element of the curriculum for financial analysts that “[a] though understanding of practical problems requires an in-depth understanding of underlying theory” (Piros & Pinto, 2013, p. xiv). In academia, theories about investment approaches have a long history reaching back at least to early documented economic thought and the theoretical ideas of Martín de Azpilcueta’s (1491-1586) *Commentary on the Resolution of Money*, first published in 1556 (Grabill, 2007). More detailed theories regarding investment valuation and equity research developed after the stock market crash in 1929 (Fox, 2009) starting from Irving Fisher (1930) and John B. Williams’ publication *The Theory of Investment Value* (1938). There have also been best-selling books such as Benjamin Graham's *The Intelligent Investor* (2003) and Burton G. Malkiel's *A Random Walk Down Wall Street* (2007) as well as comprehensive valuation guides like *Damodaran on Valuation* (Damodaran, 2006) or *“Valuation: Measuring and Managing the Value of Companies”* (Koller & McKinsey and Company, 2010). Traditional equity research approaches are either based on a fundamental analysis or a technical analysis (Damodaran, 2006; Malkiel, 2007).

2.2.1 Traditional equity analysis approaches.

“Technical analysis is essentially the making and interpretation of stock charts . . . . Charts, of course, can tell only what others players have been doing in the past” (Malkiel, 2007, p. 101) Technical analysis makes use of trend analysis and time series analysis. The basic idea is to find particular patterns or movements that might provide an indication of the further movement of a certain share price (Edwards & Magee, 1997). According to Malkiel, the technical analysis approach is used by those who believe in what he calls "castle-in-the-air theory", also known as "greater fool theory" or "survivor investing" (Keynes, 1936; Leamer, 2003; Malkiel, 2007). The castle-in-the-air view of stock pricing is largely based on psychological factors (Malkiel, 2007). In 1936, John Maynard Keynes already stated that many professional investors do not determine the proper value of an investment, but rather anticipate how the crowd of investors might act during optimistic periods, in so called bull markets, with their expectations and hopes 'castles in the air'. To be successful an investor only has to buy before the crowd builds the castle too high (Keynes, 1936).

According to Malkiel (2007), most equity analysts think that technical analysis is not a reliable tool and thus that it is somewhat unprofessional. That is one of the reasons why about 90% of the Wall Street analysts prefer fundamental analysis. Fundamental analysis is a quite different approach. It is intended to estimate the intrinsic value of an investment. There are several methods of determining the intrinsic (or firm foundation) value of an investment in place. Some of these methods are based on the assessment of the current situation of the investment; examples are the price/book value ratio, price/earnings ratio, or the sum of the parts method (Gordon, 1962; Koller & McKinsey and Company, 2010; Penman, 2007).

In practice, even more relevant for the forecasting of future share prices are the methods that take the future into account, in particular future money streams (Damodaran, 2006; Ryan, 2007). These methods are generally based on the idea that future earnings and cash flows need to be discounted in order to compare them with the investment. Financial analysts learn that "[e]quity markets respond to anticipated growth in earnings" (Piros & Pinto, 2013, p. 694). Examples are the discounted cash flow method (DCF) (I. Fisher, 1930; Ryan, 2007; Williams, 1938), the dividend discount model (Penman, 1998; Ryan, 2007), or return on equity (Koller & McKinsey and Company, 2010).

Nevertheless, it still holds true that even when the best forecasting models might deliver a good approximation of the internal or firm value of any asset, at the end of the day the existing demand and supply, which are influenced by many factors, determine the stock price (Ricardo, 1817; Smith, 1776).

2.2.2 Efficient market hypothesis and behavioural finance.

There are doubts that it is possible to outperform the markets using information. An expression of these doubts is formulated in the efficient market hypothesis (EMH) and the idea of a "random walk" of stock prices (Dupernex, 2007; Fama, 1970; Fox, 2009; Malkiel, 2007). The origins of this idea are accredited (Courtault et al., 2000; Davis & Etheridge, 2006; Fox, 2009) to Louis Bachelier's (1900) doctoral thesis *The Theory of Speculation.* However the introduction of the EMH to a wider audience was not evidenced before the 1960s. One of its early proponents was Eugene Fama. In 1965 Fama stated that an "efficient market, . . . is a market where prices at every point in time represent best estimates of intrinsic values" (Fama, 1965, p. 94). Today, the EMH is generally known in three different types: the weak, semi-strong and strong forms. According to the definition used by Fama (1970) the three types could be described as follows:

> *First, weak form . . . , in which the information set is just historical prices, are discussed. Then semi-strong form . . . , in which the concern is whether prices efficiently adjust to other information that is obviously publicly available (e. g., announcements of annual earning, stock splits, etc.) are considered. Finally strong form . . . . concerned with whether given investors or groups have monopolistic access to any information relevant for price formation are reviewed.* (p. 383)

Furthermore, Fama states that "we shall conclude that, with but a few exceptions, the efficient markets model stands up well" (p. 383). This means that according to the weak form technical analysis cannot lead to outperformance in the long run. The semi-strong form leads to the conclusion that neither technical nor fundamental analysis can generate excess returns over a long period. The strong form, as all information is already reflected in the stock-market prices, implies it is impossible to generate outperformance even with the knowledge of insider information (Beechey, Gruen, & Vickery, 2000; Dupernex, 2007). The reasoning for the EMH is that the market price perfectly reflects the relevant information, even when it is distributed among many market participants.

Contrary to the case in EMH, there are phases where the market participants sometimes seem to be irrational. In extreme forms this could even lead to mass hysteria which in turn causes “bubbles” in the market (Fox, 2009; Komáromi, 2006). “The speculative bubble is as much an error of decision-making and judgment as confusion of the inverse, hindsight bias, or the gambler's fallacy. . . . What makes the bubble more complicated, however, is the fact that it is a social phenomenon” (Freifeld, 1996).

A part of the reason for the creation of bubbles might be that in opposition to the “rationality” of the EMH, humans are not always rational (Ariely, 2009, 2010; Brafman & Brafman, 2009). A growing body of literature deals with the issue of human irrationality and markets, mainly as part of the relatively new (i.e., in academia) topic of “behavioural finance” (Akerlof, 2009; Shleifer, 2000; Thaler, 2015; Zweig, 2007).

Myriad empirical studies have been conducted to validate the EMH, but also to assess whether traditional equity research methods offer predictability in the development of share prices (Bolliger, 2004; Clement, 1999; Dupernex, 2007; Fleischer, 2005; Ho, 2012; Stanzel, 2007) and to understand analysts' behaviour and biases (Aiolfi et al., 2009; Hui, Wei, & You, 2013). According to Dupernex (2007) evidence suggests in general:

> *That markets are to a certain extent predictable. This does not mean that there are opportunities for arbitrage though, because these would soon be exploited and then*

vanish. In the real world (with taxes, transaction costs etc.) you can have some predictability without there being profitable opportunities. (p. 177)

The same is true for studies conducted in order to examine the accuracy of analysts' forecasts. The forecasts of analysts using traditional equity research methods in general deliver no advantage for the investor after transaction costs (Malkiel, 2007; Stanzel, 2007). This leads to an awareness that investment decision-making is still a very challenging task. Ho (2012) provided an indication that the quality of predictions by equity analysts may vary in different market situations and also in different countries. He concludes his thesis paper with the suggestion “that future studies should further explore the change in analyst forecast characteristics and analysts’ use of information after the financial crisis and across countries” (2012, p. 179). Inspired by his suggestion, this report focuses on the German market and analysts based in Germany.

2.3 Decision-Making and Forecasting

Decision-making and forecasting are complex processes. Benjamin Franklin used a method of decision-making by which he tried to structure the decision-making process: He suggested creating a list with two columns, one with pros and one with cons of the alternative decisions. Then he strikes out one or more of the arguments according to their relative weight. The side with arguments left is the one with the preferable alternative (Yoon & Hwang, 1995). This approach is based on the assumption that all relevant arguments are known. In complex environments, this might not be very likely. Literature suggests that there are many factors which have an impact on rational choice, decision-making and forecasting. Models such as the “Prospect Theory” (Kahneman & Tversky, 1979) are very popular in behavioural finance (Barberis & Thaler, 2003; Camerer, Loewenstein, & Rabin, 2011; Dhami, 2016). Herbert Simon’s model received less attention in the literature, but his model might be even more interesting in the context of financial markets, where incomplete information appears to be inherent to the subject matter. Simon (1955, 1956) introduced an enhanced model, the “Behavioral Model of Rational Choice”: within this model he proposed *bounded rationality* as an alternative to the “economic man” and “utility function”. Simon's model takes into consideration that access to information and the computational capacities of man are limited (Simon, 1956). Sometimes more information is also counter-productive to making a right decision. This phenomenon was demonstrated by Gerd Gigerenzer and his colleagues in their experiments with students. They asked the students questions like: Which city has more residents—Detroit or Milwaukee? The answer from students of an American college class was about 40% for Milwaukee, while the others were for Detroit. The same question answered by German students offered a clearer picture: Almost all gave the correct answer: Detroit. This is not due to the fact that German students know more about American geography than Americans; the opposite is true. They know very little about Detroit and many of them have never heard about Milwaukee. However, the German students followed a simple but successful rule while answering the question: If you know the name of one city, but not the name of the other, it is very likely that the city you have heard of has more residents (Gigerenzer, 2008). This means that more knowledge and information being available does not necessarily lead to better decisions—a finding that could be especially useful in the examination of investment decisions, where probably no one has complete information. But precise quantitative information has also its limits as “Social scientists who study the human thought processes . . . have increasingly found themselves trying to explain and overcome the paradoxical need for numbers and the numbing, desensitizing effects for quantitative disclosure” (Slovic & Slovic, 2015, p. 1). Still, it might be helpful to keep the idea in mind that forecasting “is a skill that can be cultivated” (Tetlock & Gardner, 2015, p. 4). However, it is still not an easy task to facilitate its cultivation. It might be the case that “complex models often give more precise (but *not* necessarily more accurate) answers, they can trip a forecaster's sense of overconfidence” (Silver, 2012, p. 225).

When it comes to group decision-making, there are also some special characteristics to observe. The following sections will introduce some of these characteristics, as well as the basic concepts and ideas of group decision-making.

2.3.1 Financial decision-making

Beside the different analytical techniques and methods there may also be an influence on individual decision-making within the personality of the decision maker. One example is the impact of an individual tendency to intuitive decision-making and emotions. The effect of intuitive and deliberate approaches to decision-making is a field that is of interest not only within academia, but also for business and many other fields (like politics, prosecutors). In the influential and well perceived book *Heuristics and Biases - The Psychology of Intuitive Judgement* (Gilovich, Griffin, & Kahneman, 2002) there are a few chapters that are focused to a large extent on the effects of intuition on decision-making (e. g., De Bondt & Thaler, 2002; Tversky & Kahneman, 2002). However, there are still numerous unanswered questions regarding the effects of intuitive and deliberate decision-making approaches. In the same book you can read a chapter on financial analysts' decision-making which concludes with the observation that financial analysts are not always rational and ends with the question: "After all, are not these practitioners the very same "smart money" that is supposed to keep markets rational?" (De Bondt & Thaler, 2002, p. 685). In subsequent years many researchers contributed to gaining a better understanding of the effects of intuition, conscious analysis and rationality on the decision-making and forecasting quality (e. g., Acker, 2008; Aczel, Lukacs, Komlos, Aitken, & others, 2011; Harteis & Gruber, 2008). Nevertheless, it appears that there are many uncertainties in this field. The discussion about the difference between intuitive and deliberate processes in judgement and decision-making, like dual process models and beyond, is an ongoing process in academia (Glöckner & Witteman, 2010). A particularly striking example might be the observation that the role of intuition in the process of recruiting experts and managers is an ongoing discussion in academic literature. Taneja & Arora (2015) suggest “the use of reliable and validated tests to measure managerial inventiveness” (p. 307). This thesis makes a contribution in this context, and presents some data sets and interpretations of individual decision-making behaviour from an online experiment that helps to inform our understanding of the underlying processes.

2.3.2 Group decision-making

In 1907, Francis Galton stated that “under the right circumstances, groups are remarkably intelligent, and are often smarter than the smartest people in them” (Galton, 1907, pp. 450–451). His statement raises some questions: What are the “right” circumstances and what does “often” mean exactly? To answer these questions, it is necessary to take a closer look at the group decision-making process. Some 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). Others, such as the philosopher Friedrich Nietzsche (1989) wrote that madness is rare in individuals, but he regarded it as the rule in groups, while Gustave Le Bon regarded crowds as “organisms”, but argued that they can never attain a high degree of intelligence (2009). These two counter, but equally compelling views—regarding groups as “smarter” or groups as unintelligent—rest on how the particular author views the “operation” of the group, and lead to an examination of the issues that influence such operation.

A cornerstone in the development of group decision theory was set by Condorcet (1785). He introduced what we now know as the “Condorcet jury theorem” (L. Fisher, 2009; Sunstein, 2008). In its simplest form, it states that if every group member is more than 50% likely to get the right answer, then the probability of the group reaching the right answer increases with the group size and leads to a “group intelligence”, which is a statistical result. With a probability of 60% of the individual members being right, the chance of a group of 17 members of being right is already about 80% and it is 90% for a group of 45 decision makers (Sunstein, 2008). There are also some preconditions for this theorem (Sunstein, 2008, pp. 77–78):

- the individuals in the group must be independent, which means that they must not influence one another's opinions
- they must be unbiased
- they must all be trying to answer the same question
- they must be well-informed enough to have a better than 50:50 chance of getting the right answer to the question
- there must *be* a right answer

Not all of these preconditions necessarily have to be fulfilled to arrive at good decision results. For example, if only a part of the group knows the right result and the rest of the group decides entirely randomly the majority of votes will indicate the right decision (Page, 2008b; Sunstein, 2008). An example of a simple case might be a binary decision (50:50) from a group of 30 people where 10 know the right answer and 20 decide entirely randomly. This would probably lead to 20 against 10 individual decisions which would be a clear indication. Some authors argue that there is a body of empirical and theoretical evidence indicating that there is an advantage in combining different forecasts (Armstrong, 2001; Silver, 2012). “In various disciplines, from macroeconomic forecasting to political polling, simply taking the average of everyone's forecast rather than relying on just one might reduce the forecast error, often by about 15 or 20 percent” (Silver, 2012, p. 335).

“But groups aren't perfect either. Unless they're carefully structured and given an appropriate task, groups don't automatically produce the best solution. As decades of research have demonstrated, groups have many bad habits of their own” (P. Miller, 2010, p. 59). It is just a matter of mathematics that if there is a chance of 51% that the individual decision is wrong, the probability that the group decision is wrong will increase with the size of the group (Sunstein, 2008). “In many contexts, biases and errors are systematic rather than random; in such contexts, it makes no sense at all to rely on the average answer of large populations” (p. 199). There is also a body of literature that highlights group problems and dynamics in the context of social media and online communities, e.g. the spread of misinformation, gossip dynamics, and the homogeneity of clusters where contents tend to circulate inside an echo chamber (Del Vicario et al., 2016; Quattrociocchi, Caldarelli, & Scala, 2014), “which causes reinforcement and fosters confirmation bias, segregation, and polarization” (Del Vicario et al., 2016, p. 5).

2.3.3 The social psychology of groups

According to Delbecq and de Ven (1974), “the traditional and most widely used approach for group decision-making in organizational committee life is the conventional *interaction*, or discussion group” (p. 605). In this decision-making approach, the group leader states a problem and then an unstructured group discussion and deliberation is supposed to generate ideas, exchange information among the group members, and pool opinions. “The meeting concludes with a majority voting procedure on priorities, or a consensus decision” (p. 606). Some other approaches, such as brainstorming, give the group “rules”, such as not to criticize one another during the idea generation phase (Osborn, 1963). In general, group decision-making is supposed to improve the decision-making process in some way, to avoid mistakes committed by a single person, or to legitimate a decision (Hogg & Cooper, 2003; Kahneman & Tversky, 2000; Sims, 2002). On the other hand some popular formats of group decision-making have some inherent disadvantages resulting from social psychological factors within a group (Forsyth, 1990; Sims, 2002).

Sometimes, deliberation can lead to synergy or learning, spurring creativity and producing a decision that is much better than just an aggregation of pre-existing knowledge (D. J. Cooper & Kagel, 2005). “In fact, groups sometimes do outperform their best members, in a way that suggests that synergy is involved" (Sunstein, 2008, p. 55). Much research has been done to understand group dynamics. In the 1940s Kurt Lewin started to use empirical methods and to pay attention to the prerequisites of effective group decisions (Deutsch & Krauss, 1965). Lewin also pointed out the importance of group *cohesiveness*—the positive attribution of group membership and the continued desire to belong to the group (Janis, 1982). "Lewin was most interested in the positive effects of group cohesiveness and did not investigate instances when members of cohesive groups make gross errors and fail to correct their shared misjudgments" (p. 4). Following Kurt Lewin's pioneering work more and more research has been conducted. Jenis in particular investigated the errors of judgement and faulty decisions of cohesive groups. In his book *Groupthink* he published several case studies of American foreign affairs fiascoes in order to examine group decisions. From his analysis of these cases, he came to the conclusion that "beyond all the familiar sources of human error is a powerful source of defective judgment that arises in cohesive groups—the concurrence-seeking tendency, which fosters over-optimism, lack of vigilance, and sloganistic thinking about the weaknesses and immorality of out-groups" (Janis, 1982, p. 12). He divided the symptoms of what he called "groupthink" into three main types (1982, pp. 174–175):

- Type I. Overestimation of the Group
- Type II. Closed-Mindedness
- Type III. Pressures Toward Uniformity

As a generalization from the findings in the case studies he created a theoretical model (see Figure 30 in the appendix) summarizing the antecedent conditions, the symptoms and the consequences of groupthink. This means that, apparently contrary to Galton's statement, the decisions of groups are quite often not very effective. In particular, the cohesiveness of groups might be a central factor in the assessment of the quality of group decisions.

Further research on group decisions has shown more and more difficulties and hurdles in decision-making (Diehl & Stroebe, 1990; Paulus, Dzindolet, Poletes, & Camacho, 1993). “When a group discusses an issue, it can spend too much time going over stuff everybody knows, and too little time considering facts or points known only by a few. Psychologists call this 'biased sampling'” (P. Miller, 2010, p. 59). Problems in decision making are not only found in cohesive groups, but in general research has shown that groups show a tendency to conform around particular views. This was demonstrated in a very illustrative and impressive way by Salomon Asch (1952, 1956) and his colleagues with his famous conformity experiments.

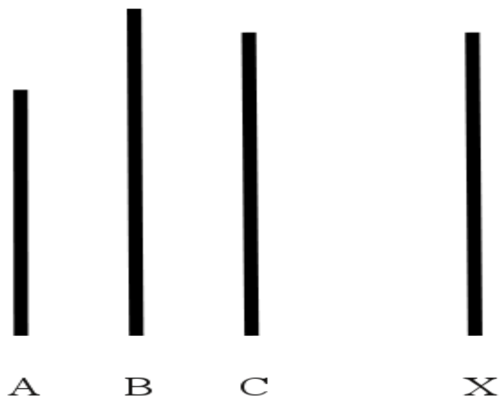


Figure 1. Illustrative example of Asch's conformity experiment: Which is the longest line?

The Asch conformity experiments produced interesting insights into group decision making and the power of conformity. The questions asked in this study were very easy to answer correctly (see Figure 1). In a control group, with no pressure to conform to an erroneous view, only 1 subject out of 35 ever gave an incorrect answer. However, when surrounded by individuals all voicing an incorrect answer, participants provided incorrect responses (X) on a high proportion of the questions (37%) (‘Asch conformity experiments’, n.d.). Three-fourths of all respondents answered wrongly to one or more questions (Asch, 1956).

“The tendency of groups to conform can be found, in particular, in face-to-face groups and is slightly more common for women” (Forsyth, 1990, p. 210). Often one of the group members will dominate the discussion.

> *Research suggests that groups whose members are familiar may be more effective at pooling information and integrating alternative perspectives than groups whose members are not familiar. Paradoxically however, the more familiar group members are with one another, the less likely they are to possess unique knowledge or differing points of view.* (Gruenfeld, Mannix, Williams, & Neale, 1996, pp. 12–13)

This means that the first stage of every decision-making process, where the major challenge is to identify the available alternatives, is very vulnerable to being undermined by the group’s behaviour. Many ideas group members might have, but not express adequately, do not even get considered in the decision-making process. It appears that an important criterion is whether the members know about the decisions of other members before they decide for themselves. This could imply that it might be important for an online group as well, if the online platform indicates the decisions of other group members before input of the individual's decision.

Some groups do not create conservative estimates and forecasts, but rather tend to develop more extreme positions. That means that a “risk-shift” occurs in the decision-making process (Wallach, Kogan, & Bem, 1962). Group members hear arguments from others that support their own position (Gigone & Hastie, 1993; Larson, Foster-Fishman, & Keys, 1994). As a result of this reassurance, the individual group members tend to a further extremization of their own position and in a next step to a greater extremization of the group as a whole. The reflection of the group’s own opinion as well as the public repetition of the group's own opinion and the arguments from others lead to a strengthening of the group's opinion (Brauer, Judd, & Gliner, 1995; Tesser, Martin, & Mendolia, 1995).

Furthermore there are problems in the process, given that group members may tend to show certain behaviour in order to gain social acceptance and avoid social hostility, through arguments that are socially desirable. This might lead to the so-called *primus inter pares* effect or the superior-conformity of self phenomenon by which group members tend in general to present themselves as "more in the norms" of the situation (Codol, 1975). Other group members might also want to avoid fitting in with the group norms, and to differentiate themselves from the others. These effects could cause a polarization of the group's norm, and/or a reduction in the variety of opinions (Hogg, Turner, & Davidson, 1990).

Although these unstructured modes of group decision making are very common, research has shown they may not be as effective as individuals working independently (Diehl & Stroebe, 1990; Mullen, Johnson, & Salas, 1991; Nijstad, Stroebe, & Lodewijkx, 2003).

2.3.4 Structured group decision making.

To overcome some of the problems of groups, there are several processes aimed at structuring decision making and reducing negative group influences. One of these processes is the nominal group technique (NGT), also called the multi-voting technique, which was designed by Delbecq and de Ven in 1968 to structure the process of decision making in order to improve decisions. NGT is designed to overcome the dominant influence of individuals in face-to-face meetings. Many variations exist, but in general the NGT proceeds as follows (van de Ven & Delbecq, 1974, p. 606):

(a) Individual members first silently and independently generate their ideas on the problem or task in writing.

(b) This period of silent writing is followed by a recorded round-robin procedure in which each group member (one at a time, in turn, around the table) presents one of his/her ideas to the group without discussion. The ideas are summarized in a terse phrase and written on a white board or the equivalent.

(c) After all individuals have presented their ideas, there is a discussion of the recorded ideas for the purposes of clarification and evaluation.

(d) The meeting concludes with silent independent voting on priorities by individuals through a rank ordering or a rating procedure, depending on the group's decision rule. The “group decision” is the pooled outcome of individual votes.

As well as the NGT, there are many other techniques in use. Another example of a structured decision-making process is the 6-3-5 Method, which is also known as the brainwriting method, developed by Bernd Rohrbach (1969). These techniques, as examples of structured decision-making processes, have in common that they try to generate more ideas by making all group members write ideas down to avoid the group members being influenced by the ideas of other group members (Brahm & Kleiner, 1996).

While solving some of the problems in the decision-making process of groups, structured group decision making, with methods like NGT or 6-3-5, could create new problems such as limited flexibility, reduced creativity and the need for preparation (Brahm & Kleiner, 1996; Sample, 1984).

2.3.5 Remote group decision making.

One possibility of remote group decision making without the use of electronic systems is the Delphi method. The Delphi methodology was developed in the 1950s by Norman Dalkey and Olaf Helmer-Hirschberg at the RAND Corporation (Dalkey & Helmer-Hirschberg, 1962). In contrast to the NGT, the Delphi method does not require the physical presence of all group members. “This approach was developed in order to reduce the shortcomings of individual thinking, opinion polls, and brainstorming” (Duckworth, Gear, & Lockett, 1977, p. 42). While many variations exist, the basic idea of the Delphi method is to gather ideas and estimates from experts by using a questionnaire. The process proceeds essentially as follows (Cuhls, n.d.; Dalkey & Helmer-Hirschberg, 1962; Fischer, 1978):

1. Define the problem and create a questionnaire.
2. Recruit people to the Delphi group and send them the questionnaire.
3. Collect the questionnaires and consolidate the answers.
4. Distribute the consolidated answers to the group.
5. Repeat steps 3 and 4, if necessary (usually 2-3 rounds).
6. Summarize all answers to create the final report.

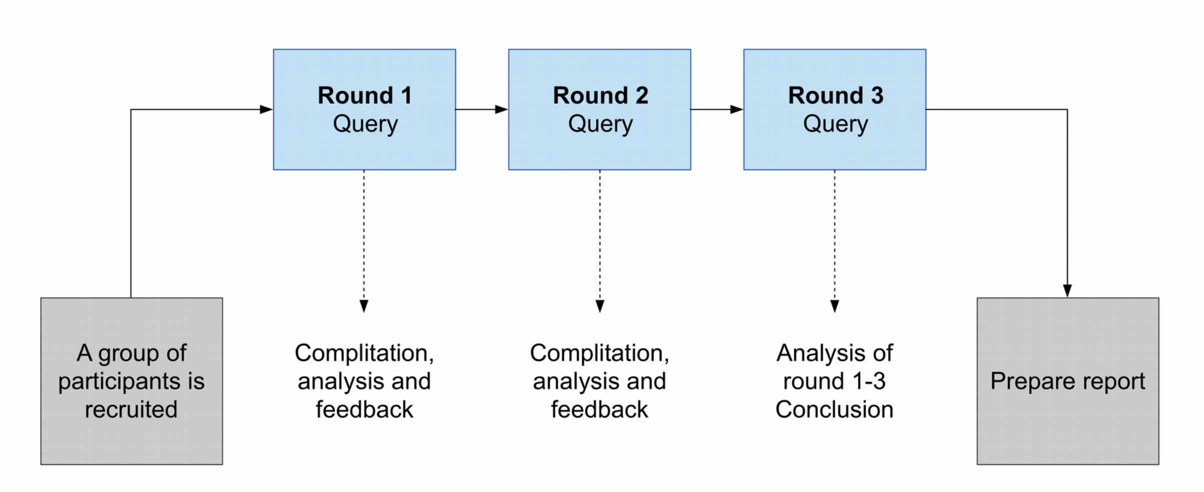


Figure 2. Basic schema of the Delphi process

Note. Adapted from (Mis- ?) using the E-Delphi Method: An Attempt to Articulate the Practical Knowledge of Teaching by P.(Source Lindqvist & Nordänger, 2007, p. 10)

The answers are collected and sent back to the participants in an anonymized form. With this information given to them, the experts in this second step are asked to rethink and adjust their answers. The adjusted answers are collected as well and a summary of the answers will again be provided. At the end all answers are analysed and consolidated in a final report. Despite its age, the application of the Delphi method can still be found in a wide range of academic fields and publications (e.g., Ballantyne, Hughes, & Bond, 2016; Pezaro & Clyne, 2015; Varho, Rikkonen, & Rasi, 2016).

Some critics of the Delphi method state “that its principal value is not as a method for predicting the future but as a method for polling large numbers of people . . . and as a heuristic device for suggesting developments” (Fischer, 1978, p. 70). Through the formal process, creativity might suffer or “opinions may not converge in the voting process, cross-fertilization of ideas may be constrained, and the process may appear to be too mechanical” (Sample, 1984, p. 1) and, performed in the original form, still require all group members to be at one place at the same time or, with some delay, by post. (An electronic and faster form of the Delphi method will be discussed later on). The Delphi method could be regarded as a facilitation framework for swarm intelligence or as O'Malley suggests “the Delphi technique has many similarities to the consensus building of bees. . . . As information is accumulated and shared, the attitudes of the group members converge on one of several possible solutions, as occurs in the hive” (2010, pp. 46–47). Since the Delphi method follows a fairly simple algorithm and is quite well documented in the literature it appears to be a reasonable starting point for further research.

2.3.6 Prediction markets

A special form of group decision making are the fairly recently emerged prediction markets (PM). PMs are a kind of trading exchange. Berg, Nelson, and Rietz (2003) define PMs as those which use “market values to make predictions about specific future events” (p. 79). The concept of utilizing market mechanisms to summarize private information widespread among many people derives from Hayek (1945). An early example of practical implementation of this forecasting mechanism is the famous Iowa Electronic Market (IEM) initiated in 1988 (Tziralis & Tatsiopoulos, 2007). The IEM is operated by the Tippie College of Business at the University of Iowa. The IEM is an on-line futures market where contract pay-offs are based on real-world events such as political outcomes, companies' earnings per share (EPS), and stock price returns (Berg & Rietz, 2006; Rietz, Forsythe, & Berg, 1997). Tziralis & Tatsiopoulos (2007) point out, as one of the conclusions from their extended literature review, that “PM research and applications will significantly increase in future” (p. 8).

The use of PMs to introduce “artificial” market mechanisms to determine prices seems logical, but is interesting as markets like stock exchanges are already in place. The assumption that PMs are more efficient in the aggregation of information dispersed among market participants than “real” stock exchanges is questionable. Analogously, one might question whether an Internet group is able to process information more efficiently and to create recommendations and price targets, than stock-markets. All examinations of decision accuracy and quality need to keep in consideration that the stock exchange participants have access to similar (or more) information to Internet groups (Z. Miller, 2010).

2.3.7 Collective intelligence

Collective intelligence (CI), also referred to as swarm intelligence, is somewhat different from group intelligence. CI spontaneously emerges from (sometimes very simple) interactions between the individuals of a group. Interactions may lead to a higher level of intelligence than any individual of the group possesses. Ants or honey bees are good examples of this approach (L. Fisher, 2009; Hofstadter, 1979; Seeley, 2010).

“A critical element in the design of this decision making system is the quorum size, for it turns out that it strongly influences the speed and accuracy of a swarm's choice” (Seeley, 2010, p. 212). However, “combining uniform perspectives only produces more of the same, while slight variation will produce slight improvement” (Tetlock & Gardner, 2015, p. 209). Scott Page (2008b) explained why diversity of the group is a key factor in decision making with his diversity prediction theorem:

Collective Error = Average Individual Error - Prediction Diversity

"The mathematical foundation for the theorem is the use of squared errors as a measure of accuracy" (Mauboussin, 2007, p. 5). Prediction diversity combines the average squared distance between the individual answers and the average guess. The average individual error combines the squared errors of how far each individual error is from the correct answer. And the collective error is the difference between the average of the individual answers and the correct answer (L. Fisher, 2009; Mauboussin, 2007; Page, 2008b). "Adding someone who predicts differently need not increase overall prediction diversity. Prediction diversity only increases if the additional person's predictions differ by more, on average, than those of other people" (Page, 2008a, p. 11). According to Page, "being different . . . [is] as important as being good" (2008b, p. 208). However, not only diversity, but also disagreement in groups may add some predictive value (Legerstee & Franses, 2015). Nonetheless, "groups don't always make good decisions either. Unless a group is properly organized, so that the face-to-face deliberations of its members result in collective reasoning that is broadly informed and deeply thoughtful, the group is apt to be a dysfunctional decision making body" (Seeley, 2010, p. 212).

2.4 Decision Support Systems (DSS)

In light of the difficulties accompanying both structured and unstructured decision-making procedures, it seems obvious to try to implement technical support systems to facilitate decision making. In fact DSS, usually interactive computer systems, have already been a topic in academia since the late 1950s (Keen & Morton, 1978). The definition of DSS has evolved over time. In the 1970s a DSS was regarded "as [a] computer based system to aid decision making" (Sol, Takkenberg, & De Vries Robbé, 1987, p. 1). Later in the 1970s, the systems became more interactive. In the 1980s, the systems included databases and models to improve and structure decision making (Sol et al., 1987). In the 1990s, the Internet started to influence DSS and at the end of the 1990s Web-based analytical applications became popular. From about 2000 Internet companies started to offer hosting and infrastructure services for decision making (Power, 2002). “More sophisticated decision portals have also been introduced that combine information portals, knowledge management, business intelligence, and communications-driven DSS in an integrated Web environment” (p. 4). DSS are still “gaining an increased popularity in various domains, including business, engineering, the military, and medicine.” (Flynn & Druzdzel, 2003, p. 3). According to Steven Alter's (1980) pioneering research there are three major characteristics of DSS:

1. DSS are supposed in particular to facilitate decision processes.
2. DSS should support but not automate decision making.
3. DSS need to adapt very quickly to altered environment variables or demands of deciders.

In general modern DSS have been developed to gather knowledge as well as to generate and evaluate decision alternatives. Nevertheless, DSS are available with various foci and are accordingly known under different types, such as group decision support systems (GDSS), computer-supported cooperative work (CSCW), group support systems (GSS), collaboration support systems (CSS), or electronic meeting systems (EMS) (Eom, 2001). “GDSS have focused on decision making/ solving problems, while CSCW provide primarily a means to communicate more efficiently. However, these two types of systems, decision making focused systems and communication-focused systems, are becoming indistinguishable” (Eom, 2001, p. 8).

The DSS field is already well covered in academia, but “around two-thirds of DSS research is empirical, a much higher proportion than general IS research. DSS empirical research is overwhelming positivist, and is more dominated by positivism than IS research in general" (Arnott & Pervan, 2005, p. 1).

2.4.1 Group decision support systems.

A GDSS can be defined as an interactive, computer-based system that aims to support a group of decision-makers to solve problems and make choices. GDSS in general is supposed to support groups in analysing problem situations and in performing group decision making tasks (DeSanctis & Gallupe, 1987; Gear, Marsh, & Sergent, 1985; Huber, 1984; Sauter, 2001). "A GDSS is a hybrid system that uses an elaborate communications infrastructure and heuristic and quantitative models to support decision-making" (Sauter, 2001, p. 1). It is also important that "the key aim of GDSS is to improve the group performance, whether it be of meeting productivity, the degree of satisfaction that is achieved and many other factors" (Davison, 2001, p. 1). Table 1 shows a typological overview of GDSS by time and place of the environment.

Table 1. *GDSS Typology by Time/Place (adapted from: DeSanctis & Gallupe, 1984)*

Same-Time / Same-Place (Most widely used GDSS- computers with projectors, voting tools)	Different-Time / Same-Place (audio/video recording, document sharing)
Same-Time / Different-Place (chat, team room, audio/video conferencing, screen sharing)	Different-Time / Different-Place (bulletin boards, Internet communities)

Practical examples of GDSS are the Claremont System, Colab System, GroupSystems, SAMM, Team Focus (Chung & Geoffrey, n.d.) and Teamworker (Gear & Read, 1993; Read & Gear, 1994). In practical use it was demonstrated that in some cases the use of GDSS “was undoubtedly useful in terms of providing a degree of structure to a complex task carried out by a large group of experts. It also enabled rapid identification of areas of strong disagreement, making it easy to prompt relevant debate” (Read & Gear, 1994, p. 250). The utilization of the Internet for GDSS in combination with community elements seems just a logical next step in the development.

2.4.2 E-Delphi.

A relatively new and specialized form of an Internet based GDSS that follows structured decision-making approach is e-Delphi. The original version of the Delphi method used regular mail to distribute the questionnaires among the participants. As e-Mail became more popular some researchers started to use this medium instead of letter mail to speed up the process of decision-making. This approach was often named “e-Delphi” or “Real-time Delphi” (Lindqvist & Nordänger, 2007). The next step in development was the use of Internet based questionnaires instead of e-Mail. Chien Chou (2002) described a prototype Web-based forecasting tool using the Delphi methodology in the context of educational research. Chou defined the “basic requirements for an e-Delphi system” (p. 234) as follows:

1. Provide a friendly interface that allows the project leader to develop and send questionnaires to panel members.
2. Provide a friendly interface that allows panel members to input data.
3. Perform calculations on panel members’ input entries.
4. Prepare individual questionnaires with multimedia presentations.
5. Help project leader determine the stability of each item in the questionnaire.
6. Allow project leader to monitor the execution of the study and to easily communicate with panel members.

As a further application, Chen and Yang (2004) used this approach for group decision making analysis in a Web environment to facilitate the complicated data collection, aggregation and analysis processes in a business context. They used a Web-based questionnaire and an Internet relay chat (IRC) technique to conduct the Delphi method over the Internet. In these examples, the practical application of e-Delphi in decision making, showed that it is less labour intensive and faster than the traditional method (Chou, 2002). “However, a dynamic Delphi survey may result in sharp changes of individual opinions and worse convergence of the collective group view when panellists are impacted by different local views. The reason may be that “local views produce uneven opinion pulls in the panel” (Liu & Yao, n.d., p. 10).

2.5 Internet Decision Making and Research Methodology

While the topic shows similarities to and is, of course, influenced by group decision making, the concrete question remains with regard to decision making through the Internet and, in particular, with regard to the forecasts created by stock trading communities. Looking at publication databases using the keywords “group decision making” and “Internet” shows that there is already a wide range of research available (see Table 2):

Table 2. *Screening by Using the Keywords “Group Decision Making” and “Internet”*

Literature Database(s)	Number of results	Results since 2000	Results since 2005	Results since 2010
EBSCOHOST				
EBSCOHOST Complete	164	153	123	79
Business Source Complete	74	70	55	34
Emerald – Journals	166	141	105	60
EThOS (British Library)	0	0	0	0
ISI Web of Knowledge (Thomson)				
ISI (Search within topic)	99	91	74	44
ISI (Search within title only)	1	1	1	1

Note. Last access: 06/05/2016.

Table 2 provides an overview of how many articles were published in 2000—the year of the founding of Marketocracy Inc. one of the first Online Communities with special focus on stock predictions, and shortly before Kaplan's paper (2001) was published—and later. The screening also shows how many articles were published after 2004, the year of the publication of Surowiecki's book *The Wisdom of Crowds*—and later, to illustrate the momentum of this topic.

The screening of the literature databases shows that many of the publications are dated 2005 or later. This is an indication that the topic is of growing interest and importance to academia, but there are still only very few articles with particular regard to investment decision making or stock market forecasts using Internet groups. In fact, only one paper describes the application of an Internet based group decision support system in an economic field, particularly in the field of macroeconomic decision making (Shen, Hu, Wang, Liu, & Zhao, 2001). Some initial efforts have also been made in researching whether twitter has predictive power for stock markets with special regard to the sentiment of the investment environment (Bollen, Mao, & Zeng, 2010; Vincent & Armstrong, 2010). There are quite a few articles that focus on different settings with crowds and online groups. A recent development in research is apparently that many authors focus enhancements on group design, like settings with smaller, smarter crowds or to set-up groups with top participants (Goldstein, McAfee, & Suri, 2014; Jose, Grushka-Cockayne, & Jr, 2014; Mannes, Soll, & Larrick, 2014). Nevertheless, the existing knowledge base gives only a rough picture of the current understanding of how online group decision processes work.

2.5.1 Stock trading communities.

There are many different stock trading communities available on the Internet (e. g. avidinfo.com, marketocracy.com, mystocks.de, sharewise.com, or tivid.com). The basic concept behind these stock trading communities is that a group decision will yield better investment outcomes than an individual's decision. Participants in these communities are financial laypersons, as well as professionals. They have various reasons for joining these communities, such as to gain a reputation, gather information for personal investment decisions, or simply for enjoyment purposes. One of the first communities in this field, founded in 2001, is Marketocracy. “[It] employs a . . . form of peering in a mutual fund that harnesses the collective intelligence of the investment community”, states Don Tapscott in his book, *Wikinomics* “It had recruited more than seventy thousand traders to manage virtual stock portfolios in a competition to become the best investors” (2006, p. 24). The best 100 investors' portfolios were used as the basis for the Marketocracy Masters 100 investment fund. In the first couple of years after launch of the fund, it consistently outperformed its benchmark (the S&P 500), but by mid 2004, that had become more difficult. The fund started underperforming the benchmark (see Figure 2) and many investors left the fund. The assets under management (AUM) reduced almost by half, from an AUM amount of nearly USD 100 million to USD 50 million. Obviously there was a problem with the investment decisions made by the community. For one thing, as the founders realized, the top investors got to know each other and discussed their investment ideas.

Figure 3. Index: Marketocracy Masters 100 and S&P 500 Index

Note. Index: Marketocracy Masters 100 and S&P 500 Index from inception date 05/11/2001 until 26/12/2015(‘BigCharts - Interactive Charting’, 2015).

Marketocracy even offered events where community members could meet and talk. It seems that this Internet community was affected by Janis' groupthink problem. “We started seeing a herd mentality emerge even among our best traders,” said Ken Kam, one of the company's founders (as cited in Howe, 2008, p. 175). Afterwards, they implemented changes to the site that made it impossible to see the trades of other members.

Another issue was that they concluded the pool of 100 members was too small and they did not use the full potential of the community diversity. Even when not the top performers, some group members could bring in some unique knowledge that might enhance the overall success of the fund (Howe, 2008). Having had bad years in 2004 and 2008, the overall performance of the fund is still under its benchmark and had only a 2-star rating (Morningstar, Inc., 2010), which means the fund is below average among funds in this class (Morningstar, Inc., 2008). In Germany a few companies started to offer “real money” investment products based on collective intelligence approaches. However, until now this might be considered a risky approach. One of the first collective intelligence investment funds in Germany, the H&A Sharewise, has already closed business. After some initial success in 2014, the fund didn't perform very well. Due to a lack of performance and subsequent outflow of funds it was apparently no longer reasonable to maintain the fund. Accordingly, they closed the H&A funds in September 2015 (Bredenbals, 2015). Other German investment funds based on a collective investment approach, the Investtor Fund (‘INVESTTOR’, 2015) and Intelligent Recommendations Global Growth Fund (Intelligent Recommendations GmbH, 2015) also seem to have a hard time: they struggle with low volumes and mediocre performance figures (Bredenbals, 2015). Another trend in the financial industry is so-called social trading. With social trading platforms it is possible to follow the strategy of others in community, which means replicating transactions by another trader in one's own portfolio. The other trader thus acts as a kind of tipster and gets some reward for his trades. Some authors consider social trading very promising (Everling &

Lempka, 2016), but the collective intelligence in this context might be limited. Social trading providers (like Ayondo, eToro, or ZuluTrade) are just supposed to identify top traders within the crowd of members of the community. However, it might be difficult for the individual investor to differentiate between a sensible investment strategy and a trader who was just a bit lucky while following a very risky strategy.

Apparently, it is not an easy task to realize consistent investment returns with collective intelligence or community investment approaches. In fact these stock trading communities represent a business approach that has barely been covered by academic research and very little literature is available about the quality of crowd sourced research (see Table 3).

Table 3. *Screening Using the Keywords "Stock Price Forecasting"*

Literature Database(s)	Number of results	Results since 2000	Results since 2005	Results since 2010
EBSCOHOST				
EBSCOHOST Complete	2605	1928	1513	732
Business Source Complete	2389	1762	1368	634
Emerald – Journals	5	4	4	1
EThOS (British Library)	0	0	0	0
ISI Web of Knowledge (Thomson)				
ISI (Search within topic)	76	73	67	41
ISI (Search within title only)	39	39	35	21

Note. Last access: 06/05/2016.

A screening of the literature databases indicates that most of the publications are dated 2005 or later. This is an indication that the topic is of growing interest and importance for academia. Despite this growing interest, there is still only one article with particular regard to investment decision making or stock market forecasts using an online approach. In 2001, Craig Kaplan presented a paper: "Collective Intelligence: A New Approach to

Stock Price Forecasting” (2001) at the IEEE International Conference on Systems, Man, and Cybernetics. In that paper, he described the design and first tests of a prototype CI system that is supposed to create stock trading recommendations based on input from the crowd. During his test, the system outperformed the benchmarks (in the form of market indices). He claims that “there is a growing body of evidence that the key to forecasting the stock market lies neither in value analysis nor in technical analysis. Rather, investor psychology seems to be the critical factor” (Kaplan, 2001, p. 1). In his test, performance improved as more people participated; however, his tests were still conducted with a quite small group (62 people) and only over a period of 11 trading days. He suggests that the next steps should include “conducting a test on a much larger scale, and experimenting with variations of the CI processing algorithms to identify those that are most effective” (Kaplan, 2001, p. 6). This opens a field for further research to clarify which conditions might influence the predictions and forecasts of a remote group.

2.5.2 Research framework and experimental research methodology.

Many methodologies and techniques have been developed in order to enhance the efficiency and effectiveness of group decisions. But many of these approaches have not been assessed thoroughly. Some authors have already addressed the need for a generally agreed framework within which research can be conducted and results determined (Fjerrnestad & Hiltz, 2002; Shaw, Eden, Ackermann, & School, 2002; Stevens & Finlay, 1996). Finlay and Stevens propose such a research framework, involving the identification of the context, process, and outcome variables. They suggest that these variables are likely to be important for understanding, and subsequently predicting, the appropriate forms of intervention in the workings of groups. They have highlighted some major components: the organizational environment, the group context, the process context, the group process, the substantive outcomes, and process performance indicators. They regard their framework as applicable to a very wide range of group support systems which might be used in many contexts (Stevens & Finlay, 1996).

Fjermestad and Hiltz (1998) conducted an extensive literature review of publications that examined processes and outcomes in computer-supported group decision making. They gave an “overview of what has been studied and how: the systems, independent, intervening, adaptation, and dependent variables, manipulated or measured, and experimental procedures employed” (p. 2). “If researchers learn the lessons summarized in this paper in terms of what is already known and what experimental procedures need to be followed and reported to obtain results that will contribute substantially to the field, the next generation of experiments will be very rewarding” (Fjermestad & Hiltz, 1998, pp. 48–49).

Moreover, Nunamaker published several suggestions for further research including more studies in distributed settings and virtual environments (1997).

2.6 Analysis of Existing Online Communities and Published Analyst Recommendations

During both the pilot and the main experiment, secondary data was gathered in existing online communities with a focus on stock price predictions (see Appendix: Stock Trading Communities, p. 284) as well as relevant recommendations from financial analysts for the stocks covered with the experiment design.

2.6.1 Analysis of existing online communities parallel with the pilot experiment.

The parallel with the pilot experiment was an analysis of existing online communities, which made visible some possible difficulties in the group decision-making process. In general, a lack of group member activity was observed to be problematic. In the American-based online communities examined, predictwallsteet.com and tivid.com, only the skeleton for the four shares existed, and community members made no comments or predictions. The three German communities analysed, Sharewise.com, stockjaeger.de and Spekunauten.de, had at least some comments and predictions. Still, the limited activity of the members and the time lag between the predictions might be problematic. The communities stockjaeger.de and Spekunauten.de had almost no activity and new posts during the examination period. Activity was measured in Sharewise.com (see also Figure 32 in the appendix), but the fact that recommendations sometimes stay within the community consensus for half a year before they are excluded may still pose a problem. It is possible that members change or cancel the recommendation, but it is also possible that they just post a comment and recommendation and don’t adjust if anything, e.g., market conditions or company perspectives, has changed. Possibly because of the lack of updates, there seemed to be a tendency to make very positive predictions.

Another set of secondary data for the pilot run was collected from published equity analysts' recommendations. One of the most important data vendors in the financial industry is Bloomberg. The Bloomberg Professional terminals are a very common information device for financial analysts and professional investors. As a data vendor Bloomberg also provides a service as an aggregator of analysts' recommendations. These aggregated analysts' recommendations are distributed within the financial industry as so called “Bloomberg Consensus” or Bloomberg Estimates (BEst) (see Figure 4).

Figure 4: Bloomberg Professional Terminal - BEst TKA GY Equity

It turned out that existing online communities (in this case, stockjaeger.de, Spekunauten.de, and Sharewise.com) in direct comparison, also had a very low predictive accuracy (see Figure Fehler: Referenz nicht gefunden). The group of professional financial analysts, as represented in the Bloomberg Consensus, was better than these Internet groups, but still not as good as the lay group in this pilot experiment.

2.6.2 Analysis of existing online communities parallel with the main experiment.

The parallel with the main experiment was an analysis of existing online communities with a focus on stock price predictions (see also appendix: Stock Trading Communities (Main Experiment) p. 287) as well as equity analyst recommendations available on Bloomberg Professional, the so called Bloomberg Consensus Estimates, and analyst recommendations available on the Sharewise community website.

The data analysis from the main run of the main experiment confirmed a few possible difficulties in the group decision-making process, as had already been found with the pilot run. Again, a lack of group member activity was observed to be problematic. There were still (as per 17/11/2012) no usable data in the American-based online communities examined, predictwallsteet.com, valuelessforum.com (a new community; included only in the main experiment), and tivid.com: only the skeleton for the five companies existed, and community members did not leave any comments or predictions. However, the three German communities, analysed Sharewise.com, stockjaeger.de and Spekunauten.de, had at least some comments and predictions (see Table 4).

Table 4. *Online Community Recommendations (Secondary Data)*

		Number of Recommendations 16/11/2012	Number of Recommendations 04/02/2013
Sharewise (Group)	Adidas	11	15
	Heidelberg	7	8
	RWE	11	14
	Siemens	27	20
	ThyssenKrupp	30	33
Speckunauten	Adidas	32	33
	Heidelberg	10	10
	RWE	32	32
	Siemens	41	44
	ThyssenKrupp	27	28
Stockjaeger	Adidas	2	2
	Heidelberg	3	3
	RWE	2	2
	Siemens	3	3
	ThyssenKrupp	2	2

There were significant differences in the predictive accuracy of the groups (Chi-square: 37.385, DF=2, *p-value*<0.001). However, the limited activity of the members and the time lag between the predictions may limit the informative value of the data. Again, the communities stockjaeger.de and Spekunauten.de had almost no activity in terms of recommendation changes and new posts during the examination period. As was the case with the pilot activity was also primarily measured in Sharewise.com, but the fact that recommendations sometimes remain within the community consensus for half a year before being excluded may still pose a problem. The community software offers the possibility for members to change or cancel the recommendation, but it is also possible that they just post a comment and recommendation and don't adjust if anything, e.g., market conditions or company perspectives, has changed. Possibly because of the lack of updates, there seemed to be a tendency to make very positive predictions in the main run.

Table 5. *Predictions from External Communities*

		Sharewise.com	Stockjaeger.de	Spekunauten.de
3 Month	Correct	26	61	64
	Wrong	58	19	36
Predictions	Excluded	16	20	0
% Correct	Correct	31%	69%	64%
	Wrong	69%	31%	36%

In contrast to the pilot run, it turned out that existing online communities (in this case in particular stockjaeger.de and Spekunauten.de) in direct comparison, also had quite high predictive accuracy (see Table 5 and Figure 5).

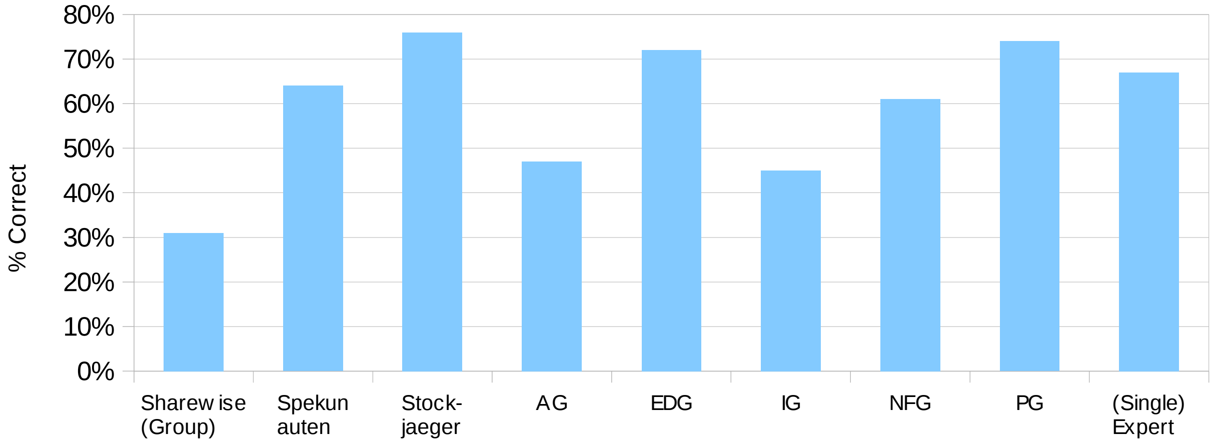


Figure 5: Group comparison: 3-month performance external Online-Communities

2.6.3 Analysis of published analysts' recommendations.

The data from financial analysts is usually acquired from several data providers, as well as by banks' internal distribution mechanism. One of the most important data vendors in the financial industry is Bloomberg. The Bloomberg Professional terminals are a very common information device for financial analysts and professional investors. As a data vendor Bloomberg also provides a service as an aggregator of analysts' recommendations. These aggregated recommendations are distributed within the financial industry as so called “Bloomberg Consensus” or Bloomberg Estimates (BEst). The Bloomberg Consensus comprises recommendations for the companies in the main experiment from about 40 analysts (see also Table 6). All stocks have been selected from five different companies in five different sectors: consumer goods (Adidas, Bloomberg code: ADS GY Equity), construction (HeidelbergCement, Bloomberg code: HEI GY Equity), utilities (RWE, Bloomberg code: RWE GY Equity), technology (Siemens, Bloomberg code: SIE GY Equity) and industry (ThyssenKrupp, Bloomberg code: TKA GY Equity). Since all companies from the experiment were selected from the German main stock index DAX the analysts' coverage was accordingly high. The analysts' coverage of a company typically depends, among several factors, to a large extent on the importance of the company for the stock market.

In the Sharewise portal there are also overviews where analysts' recommendations are available. Even though some analysts' recommendations are on Bloomberg as well as on Sharewise, it appears that compared with BEst the overview is less comprehensive. While on Bloomberg about 40 analysts' recommendations are published, there are on average fewer than 20 analysts' recommendations available on Sharewise.

Table 6. *Bloomberg (BEst) and Sharewise Analyst Recommendations*

		Number of Recommendations 16/11/2012	Number of Recommendations 04/02/2013
Bloomberg (BEst)	Adidas	41	40
	Heidelberg	38	38
	RWE	37	39
	Siemens	40	42
	ThyssenKrupp	34	35
Sharewise (Analysts)	Adidas	16	15
	Heidelberg	18	16
	RWE	21	18
	Siemens	19	19
	ThyssenKrupp	17	18

Despite the great difference in the number of recommendations, the difference in predictive accuracy is only weakly significant (Chi-square: 3.429, DF=1, *p-value*=0.064). During the main experiment the predictive accuracy on Bloomberg (74% correct predictions) was higher than the corresponding predictions on Sharewise (69% correct predictions).

Table 7. *Consensus Predictions from Financial Analysts*

		Bloomberg Consensus	Sharewise Anaysts
3 Month Predictions	Correct	76	64
	Wrong	24	36
	Excluded	0	0
% Correct	Correct	76%	69%
	Wrong	24%	31%

The group of professional financial analysts, as represented in the Bloomberg Consensus, was better than the existing Internet groups (see Table 5 and 7) and also slightly better than the purposefully composed groups from the main run of the online experiment (see Figure 6).

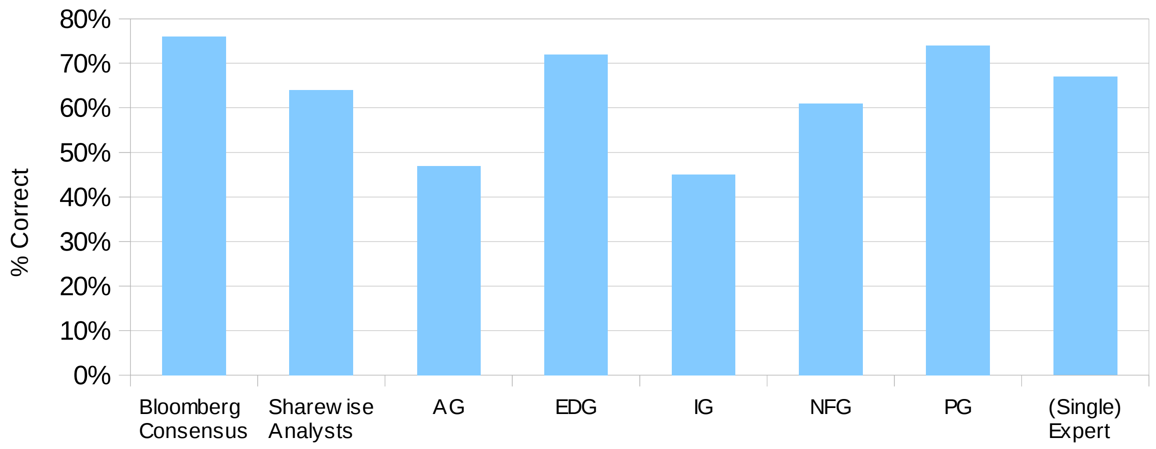


Figure 6: Group comparison: 3-month Performance Analysts' Recommendations

2.7 Conclusion of the Literature Review

In summary, the literature suggests that there is still a lack of reliable mechanisms to identify and assess investment ideas. Literature on traditional approaches for equity research highlights weaknesses and generally describes the quality of the predictions as problematic (see also 2.2 Equity Research on page 27). While there are numerous papers that focus on collective intelligence, an approach which is used by numerous Internet groups, there is very limited literature with a focus on collective intelligence approaches in equity research.

One of the key papers in this context was provided by Craig Kaplan (2001). He suggests “conducting a test on a much larger scale, and experimenting with variations of the CI processing” (2001, p. 6). The synopsis of the literature presented inspired the overall aim of the study:

to explore, analyse and compare the quality of equity predictions of individuals and groups who are using the Internet in order to build a theory of process.

Research question 1 picks up on his suggestion, and adds external benchmarking with financial analysts.

RQ1: How do the recommendations of an Internet group perform in comparison to the recommendations of an individual expert (financial analyst)?

A widely used and well-documented approach for group decision making is the Delphi process (Dalkey & Helmer-Hirschberg, 1962). This approach can be seen as a variation of a CI process as suggested by Kaplan (2001). Still, the Delphi process is not without critics (Fischer, 1978; Linstone & Turoff, 2002; van de Ven & Delbecq, 1974). Many studies using the Delphi method have no stringent follow-up, and it is often unclear whether the predictions made using the Delphi panel turn out correct or not (e.g., Cole, 2008; Hsu, 2005; Kuhn, 2004). This research project includes a follow-up on the prediction quality. Research question 2 addresses this gap in the domain of equity predictions with Delphi.

RQ2: How does the feedback loop of an e-Delphi process affect the prediction quality of an Internet group making equity predictions?

There are various concepts in literature that provide an explanatory framework for decision-making processes. While general concepts, including bounded rationality (Simon, 1955) and prospect theory (Kahneman & Tversky, 1979), outline several underlying mechanisms and allow the assessment and discussion of various aspects in decision making, they are still missing domain specific aspects. Kahneman points out that “[a] general framework . . . is not a substitute for domain-specific concepts and theories” (2003, p. 717). Research question 3 aims to contribute with domain specific insights on the underlying key mechanisms in the context of investment decisions and equity predictions.

RQ3: What are the underlying key mechanisms, of the individual and of the group, that influence the decision-making process, and how might the decision-making process in existing online communities be improved?

3 Research Methodology and Methods

This methodology and methods section is intended to outline the general structure of the research project to assess the quality of equity research in Internet communities. The aim is to present the approach in general, and methods in brief.

3.1 Research Philosophy

The research philosophy of pragmatism and a realist point of view is common in mixed-methods research (Creswell, 2009; Johnson & Onwuegbuzie, 2004). This philosophical position is supposed to address the need to conduct research within the complex process of equity research as a group decision in Internet communities in an appropriate way by using mixed methods (Creswell, 2009). The analysis of how well group decisions compare with traditional equity research and actual market results follows a positivist approach. In business research the traditional way to conduct research follows methods borrowed from the natural sciences (Patton, 1990) and much of the business research can be attributed to the positivist research paradigm (Eriksson & Kovalainen, 2008; Patton, 1990). Post-empiricist and critical theory schools also "had considerable influence upon research in financial disciplines" (Ryan, Scapens, & Theobald, 2002, p. 30). Quantitative methods are appropriate to addressing questions 1 and 2. To understand the process it is appropriate to use qualitative methods to address question 3. The results from these were triangulated to verify one against the other (Creswell, 2009), and build an explanatory theory. Qualitative methods are used to help to interpret and understand the quantitative results.

3.2 Research Questions and Objectives

As stated in the introduction, the basic motivation for this study is to address the question: 'under what circumstances would a remote group like an Internet community outperform the equity research forecast accuracy of an individual financial analyst'? Many special interest communities focus on decision-making, using a remote group process to create equity price predictions, but the literature review suggested that so far no academic evaluation of when or if this practice is effective has been conducted. This study assesses the practice in terms of the conditions which may enable it to outperform equity research experts. The aim is to develop an explanatory schema and create a theory to begin to understand why and when it happens.

To achieve this target the general research question needs to be split into two, intended to help in assessing and describing the process as well as the major input factors. Due to the complexity inherent in the group decision-making process it is not possible to examine all possible influencing factors and variables. The literature review (see also 2.7 Conclusion of the Literature Review) suggests the following specific research questions to address:

- Research question 1:

  How do the recommendations of an Internet group perform in comparison with the recommendations of an individual expert (financial analyst)?

- Research question 2:

  How does the feedback loop of an e-Delphi process effect the prediction quality of an Internet group making equity predictions?

- Research question 3:

  What are the underlying key mechanisms, of the individual and of the group, that influence the decision-making process, and how might the decision-making process in existing online communities be improved?

Obviously, there are many factors that might influence the quality of decision outcomes. The research questions imply that there might be quantitative and qualitative factors that have a major influence on the quality of the group decision. Research questions 1 and 2 are mainly addressed using quantitative methods. Research question 3 is mainly addressed using qualitative methods. However, the overall analysis and synopsis is informed by both methods and uses triangulation of both where appropriate.

3.2 Discussion and Selection of Appropriate Research Methods

In a DBA research journey, unlike a traditional PhD which usually addresses a purely academic question, the research conducted in general is supposed to deal with the academic perspective as well as concrete application in professional practice as in the context of high level strategic business issues and problems (University of Gloucestershire, n.d.; University of Southampton, n.d.). Accordingly, the knowledge production approach needs to be adjusted appropriately. Traditional approaches, in a sort of Humboldtian and Newtonian tradition of conducting research, “tends to be description-driven and is problem-focused rather than solution-focused, more interested in analysis than in design” (van Aken, 2001, p. 5). Gibbons et al used the term “mode 1” (1994) to describe this kind of knowledge production. While this approach might be suitable for creating fundamental knowledge, another approach is needed to create the applied knowledge needed for conducting a DBA. Gibbons et al introduced the term “mode 2” for this new kind of knowledge production. Unlike mode 1 Aken describes that “in contrast, mode 2 knowledge production is solution-focused, oriented not only on analyses of problems but also on designing solutions. It is often trans-disciplinary in nature” (van Aken, 2001, p. 4). Unlike mode 1, which is usually executed within academia, mode 2 knowledge production is “characterized by a constant flow back and forth between the fundamental and the applied, between the theoretical and the practical” (Gibbons et al., 1994, p. 19). While mode 2 is widely used, it is not the only “approach to study changes in science system” (Hessels & van Lente, 2008). Hessels and van Lente (2008) give a brief introduction to these approaches. One of the common characteristics is that most of these knowledge-production approaches include elements of “interaction with other societal 'spheres' (industry, government)” (Hessels & van Lente, 2008, p. 744). These tendencies to trans-disciplinary approaches and interaction with other societal 'spheres' might also be a reason for the development of new research methods. In particular, research methods that enable practitioners to contribute to knowledge production. A popular example of these new methods is action research (Anderson & Herr, 2005; McNiff & Whitehead, 2009), even though some see in action research neither a “method nor a technique” (Greenwood, 2007, p. 131). Another popular method which can be used for the generation of mode 2 knowledge is the case study method (Garvin, 2003; Thomas, 2011). Research in the context of group decision-making is a complex process. In order to accommodate this complexity, the action research approach addresses several issues. Nevertheless, the action research approach might not be the best choice for this kind of research project. One of the practical problems might be that to facilitate an online group process from which a single plan of action emerges, different participants might have different ideas about the changes needed in order to improve the predictions. Another issue is the fact that it takes time to analyse whether the action implemented solved the problem in terms of generating outperformance to the market. So the action research typical cycles might not work properly in this context.

This suggests the conclusion that major adjustments and/or enhancements to the action research approach are needed prior to the study. In order to avoid these necessary methodical preparations to adjust the action research approach, it might be more opportune to use a different approach for this research project and to answer the initial question: under what circumstances, including the mechanisms driving the decision-making process, would a remote group like an Internet community outperform the equity research forecast accuracy of an individual financial analyst? A more traditional mixed-methods approach in the form of a sequential study might be more appropriate to answer this question. An evaluation project in the form of a sequential mixed-methods evaluation project was conducted.

Some authors used simulators and controlled laboratory experiments for the examination and analysis of investment decisions (Aramburo, Acevedo, & Morales, 2009; Ball & Wingender, 1988; M. A. Bradbury, Hens, & Zeisberger, 2014),, despite the general criticism that the artificial environment of the laboratory "tell[s] us very little about how respondents would actually act in real life" (Thompson, 2016). Additionally, a simulator might be not adequate for the assessment of the underlying mechanisms of the process, as aimed at in particular with research question 3. While a simulator experiment would be essentially focused on measuring the effect of known variables, this research project also aims to identify new mechanisms and variables. The evaluation project utilized an e-Delphi approach to generate primary data. Additionally, all participants in the experiment—lay people as well as financial analysts—were interviewed. The interviews were semi-structured, and the data gathered by the experiment were evaluated to aid in the preparation of the interviews. The interviews were designed with the aim of gaining an understanding of the processes individuals used to make the decisions recorded during the experiment.

3.4 Research Purpose and Design

The purpose of this study is to begin to understand the group decision-making process for Internet communities which focus on stock-trading. The intent of this two-phase, sequential mixed-methods study is to develop an explanatory model of the group decision-making process of Internet communities. Figure 7 shows a sequential study approach to the study, with a qualitative phase building on and helping to explain the initial quantitative phase (Creswell, 2009).

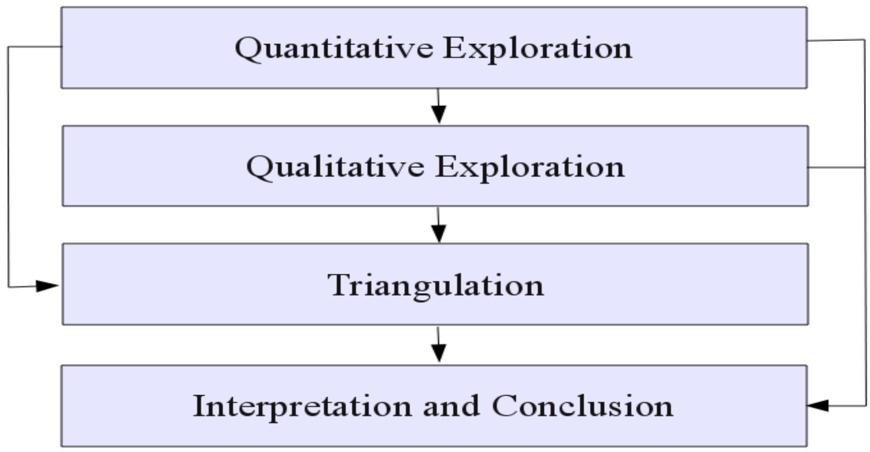


Figure 7. Sequential study approach

3.4.1 Quantitative approach.

A quantitative approach is the more traditional way to conduct business research (Bryman & Bell, 2003; Eriksson & Kovalainen, 2008; Patton, 1990). “The experimental method is the only method of research that can truly test hypotheses concerning cause-effect relationships” (Gray & Diehl, 1992, p. 382). In the first phase, quantitative research was used to address the relationship between the predictions and actual outcomes. The benchmark for these community predictions is a comparison of the group and estimates of financial analysts, with actual market results. These comparisons are aimed at measuring whether the group decision-making process is better, or worse, than the predictions of financial analysts.

3.4.2 Qualitative approach.

In the second phase, qualitative interviews and observations are used to explain the results from the first quantitative phase (Creswell, 2009; Eriksson & Kovalainen, 2008). Qualitative research has become more widely accepted during the last 10-20 years, even though the origins of methods reach far back into history. Aristotle (384-322 BCE) is sometimes referred to as the founder of qualitative research (Mayring, 2002). The reason for following up with qualitative research in the second phase is to better understand and explain the quantitative results of phase 1.

3.4.3 Mixed methods.

The combination of quantitative and qualitative approaches and the triangulation of both promises to create a more holistic understanding of the decision-making process of these communities associated with a "pragmatic perspective where designs and methods are selected on "what works" for answering the stated research questions" (Plano & Badiee, 2010, p. 279). Where the research questions consist of confirmatory and explanatory questions, mixed methods become appropriate (Teddlie & Tashakkori, 2009).

3.5 Research Data

The research project is based on empirical data and appropriate methods (Bortz & Döring, 2015). The data for this research was gathered primarily from two sources: a controlled experiment and interviews. These data are supplemented by data from existing stock-trading communities, financial data providers (like Bloomberg, Thomson Reuters or Yahoo-Finance), books, journal articles, newspaper stories, miscellaneous papers and documents.

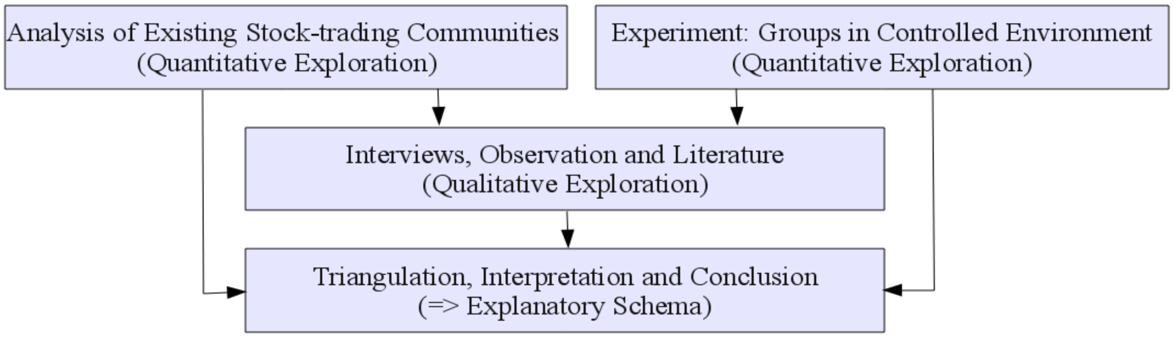


Figure 8. Research design and data generation

3.5.1 e-Delphi experiment to generate primary data.

In existing stock trading communities on the Internet the data availability and quality might not be reliable (see also 2.6 Analysis of Existing Online Communities and Published Analyst Recommendations). For this reason, in addition to the collection of secondary data, a field-based experiment was conducted in a defined and controlled environment. Financial analysts and a group of lay people using a defined process (e.g. e-Delphi) generated the data needed. The field-experiment was conducted following an e-Delphi approach. Every e-Delphi cycle in this experiment consisted of a first round for data collection. This data was compiled and distributed among the panel. In a second round, participants are allowed to give different answers in respect of the feedback they got from the group's decision in the first round. This empirical data is *primary data* (Bortz & Döring, 2015; D. Cooper & Schindler, 2011; Eriksson & Kovalainen, 2008), collected purposefully for this research. Data gathered from other sources like existing communities or data vendors is called *secondary data* (D. Cooper & Schindler, 2011; Eriksson & Kovalainen, 2008). A quantitative exploration of these data sources may help to address research questions 1 and 2. An assessment of the basis of individual decision-making by group members, the accuracy of each member's individual decision, the learning effect through the feedback loop, and the quality of the group's decisions were examined using the data from the experiment.

The mainly quantitative assessment is intended to answer the first two research questions. The qualitative exploration using the data gathered from the quantitative analysis is intended to address research question 3 (Eriksson & Kovalainen, 2008). All participants in the experiment–lay people as well as financial analysts–were interviewed. The interviews were semi-structured and the data gathered in the experiment was evaluated to prepare the interviews. The interviews were designed with the aim of gaining an understanding of the process used by individuals to make the decisions recorded during the experiment.

3.5.2 Secondary data as benchmark for the e-Delphi experiment.

In order to measure the quality of the decisions made by these communities, a benchmark is needed. One benchmark is actual market development. Market data is made publicly available by stock exchanges or via several data vendors like Bloomberg, FactSet, Yahoo-Finance or Thomson Reuters. Another benchmark is the data generated from the predictions of financial analysts. These predictions are publicly available as single analyst predictions and estimates or aggregated as the so-called analyst consensus. The consensus data is an average of the estimates by financial analysts provided by data providers like Bloomberg (BEst) or Thomson Reuters (I/B/E/S). The consensus data utilized as a benchmark in this study is based on the Bloomberg consensus data.

By an additional examination of existing stock-trading communities on the Internet a large pool of secondary data was purposefully utilized for this research. The data needed for benchmarking the results from the e-Delphi experiment was gathered to a great extent from these communities. In particular, this took place with the special interest stock-trading communities Sharewise.com, Spekunauten.de, stockjaeger, predictwallsteet.com and valuelessforum.com.

3.5.3 Ethical standards.

It is also necessary to consider that some experimental set-ups are not possible or, due to ethical standards and regulations, not acceptable any more. For example the study of the effects of group pressure by Stanley Milgram (1964) would not be suitable nowadays. The fact that some of the participants in this study might suffer after the experiment is not acceptable (Gray & Diehl, 1992). With the planned experiment no one is likely to suffer physical stress, but it might be difficult to deal with the issues related to predicting share prices. It is probable that these predictions will be wrong (sometimes) and participants may be afraid of negative consequences like a loss of reputation, i.e. 'evaluation apprehension' (Bordens & Horowitz, 2001). The experiment could also interfere with the protection of proprietary information from companies. These issues are particularly relevant for professional financial analysts. In order to avoid these possibly negative consequences the anonymity of all participants was ensured and made clear, and agreed by participants, at all stages of the planned experiment. Signed consent to participation was obtained from each group member well in advance of the experiment.

3.6 Sampling

Deciding on the appropriate sample size for the research is not easy. One simple answer is “large enough!” (Gray & Diehl, 1992, p. 140), but the authors who gave this answer admitted that it is “not very comforting” (Gray & Diehl, 1992, p. 140). In particular for the Delphi method, “the average error of the group responses declined monotonically with the size of the group, with decreasing returns with increasing size” (Dalkey, n.d., p. 1). According to Dalkey, Brown and Cochran (1969) the minimum size of a group for a Delphi process is not sharply defined. They created a curve to show the effect of group size (see Figure 9) and “selected 7 as the lower limit on the grounds that it was roughly in the middle of the “knee” of the curve” (Dalkey et al., 1969, p. 6). However, this is now very old data and analysis in a different contextual task.

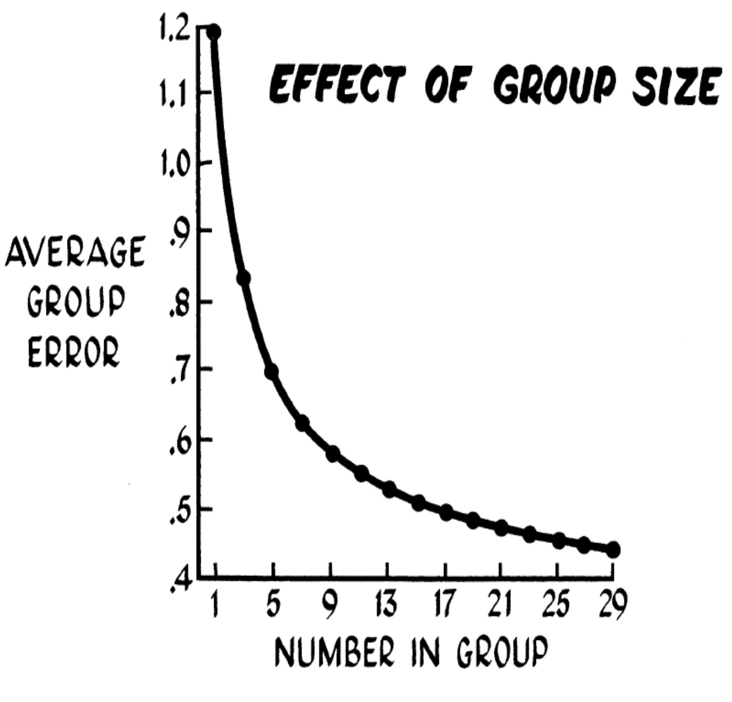


Figure 9. Effect in group size

Note. Adapted from “The Delphi method : An experimental study of group opinion,” N. C. Dalkey, 1969, *Rand Corporation*, p. 11.

Another indication that might help to determine the group size in the context of CI comes from Ashley Ward and colleagues who obtained similar results from an observation of fish shoals: a group of 8 fish is considerably better at avoiding predators than a group of four, but there is no great difference (see Figure 29 in the appendix) when the group size increases from eight to 16 (Ward, Herbert-Read, Sumpter, & Krause, 2011). These statements suggest using a group size not lower than eight participants in each round. Since it is also necessary to keep in touch with all participants over a period of three months and to allow for a moderate drop-out quote an initial number for the panel of about 10 participants (*N*=10) appears appropriate for the lay person group in the pilot study. The pilot run generally confirmed that it is possible to handle an experiment with this sample size. However, to compare different survey designs and measure differences more than one lay group was included in the main experiment. In addition to the determination of an appropriate sample size, it is necessary to consider possible sampling bias. This might not guarantee that there are no errors in sampling, but should help to avoid systematic bias where possible (Gray & Diehl, 1992). In order to meet these concerns, all participants in the experiment were purposefully selected from the personal and professional network of the researcher. A key criterion for selection is ensuring the diversification of the group in terms of age, gender, education-level and professional background, etc.

Additionally, ten independent “cycles” of e-Delphi were conducted over a period of ten weeks and four different shares from different sectors were assessed in order to cover several market patterns. In order to ensure that different market situations are covered these ten e-Delphi cycles have been separated from each other by about one week. The participants are asked at these ten different points in time to provide their estimate, each time one question round with one feedback loop that allows revising or confirming the first answer. This meant that the overall data collection period of this field-experiment took about three months (see Figure 11).

The participants are acquired from the extended personal network of the researcher to provide a purposeful sampling of different, but comparable, groups of laypeople and financial professionals. In the pilot run financial professionals were only represented by equity analysts, but in the main run a second group of investment professionals (equity trader, portfolio manager, etc.) was included as well. All the expert participants in he main experiment, financial analysts (AG) and professional investors (PG); ten professionals in all, were from four different financial services companies with offices in Germany. All individuals in the professional groups were highly qualified and had access to several professional investment information services (e.g., Bloomberg, Thomson Reuters, industry reports, in-house research material). The first group of professionals, the analysts, consisted of five financial analysts with many years of industry experience. All the forecasts by the analysts are included in the group results of the analysts' group, including forecasts for stocks within and outside their professional coverage. Financial professionals who were not investment experts like regular bank clerks, insurance brokers etc., were excluded from the sampling.

Figure 10 shows the experimental design of the research. The primary data collection follows an adapted e-Delphi method. All participants are asked to complete an online questionnaire every in regular cycles. The experiment consisted of several e-Delphi rounds and an e-Delphi round of 2 queries. Every Friday the first query of a round is open. The participants receive an individualised link to open a questionnaire based on an online form. Depending on the respective group design of the assigned group of the participants, the results from the group are compiled and distributed back to the group, anonymised and on aggregated level. However, there are also group designs without feedback loop and group designs with interaction between the group members. These differences in group design are intended to allow assessment of the impact of the feedback loop on the quality of the predictions. On the following Monday the second query of each e-Delphi round was conducted. The second round allowed the participants to adjust or change their recommendations in the light of the group feedback.

After each round the final results are also compiled and—depending on the respective group design—distributed back to the participants. Since the equity predictions in this experimental design are based on existing listed companies there are also news flow and market development during the experiment. Participants are allowed to use any information available to them for the experiment. The online questionnaire as well as the accompanying in-depth interviews include questions aiming to identify patterns of information influence on the decision-making process of the individuals and groups.

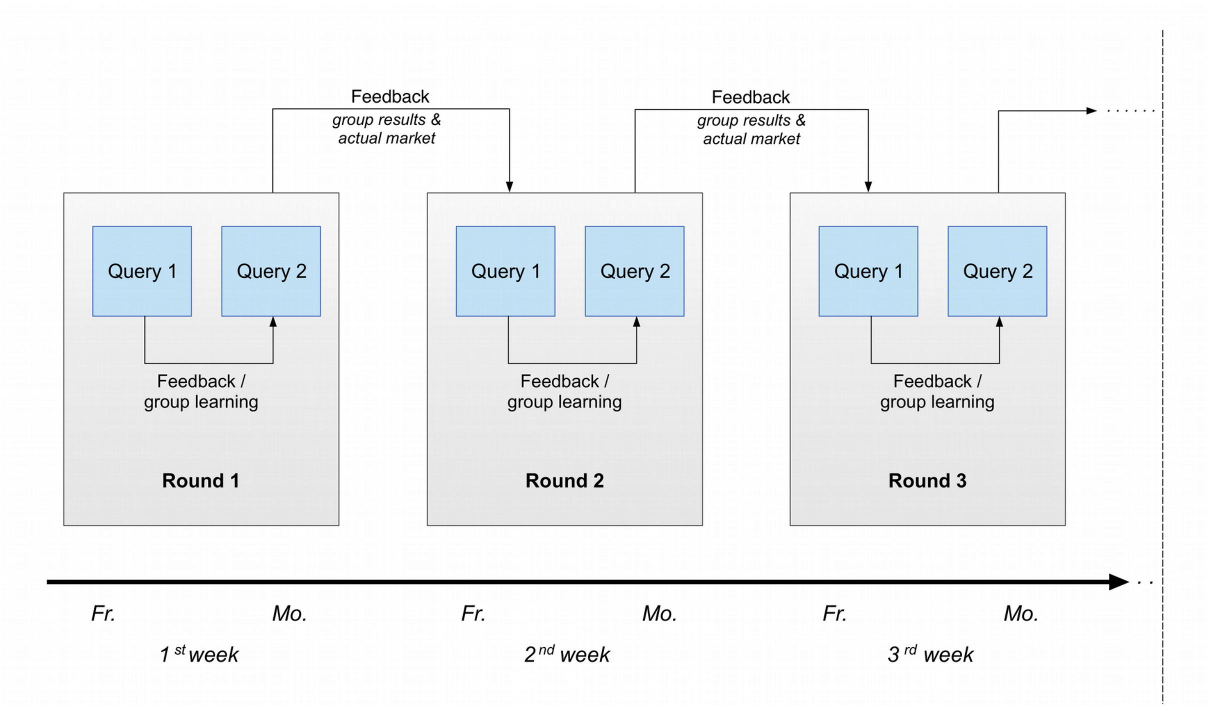


Figure 10. Primary data collection with e-Delphi method

It might be the case that in different market situations certain decision-making approaches are more likely to generate correct recommendations. Therefore, the results of these three e-Delphi cycles were compared against each other to indicate whether good predictions might be more likely in a particular market environment.

As another benchmark for comparing the decision results of the group of lay people, financial analysts need to take part as well. Ideally there are three to five analysts participating in order to have at least one analyst to cover each respective share. This means that the selection of shares was largely defined by the coverage universe of the participating financial analysts.

3.7 Methods and Procedures

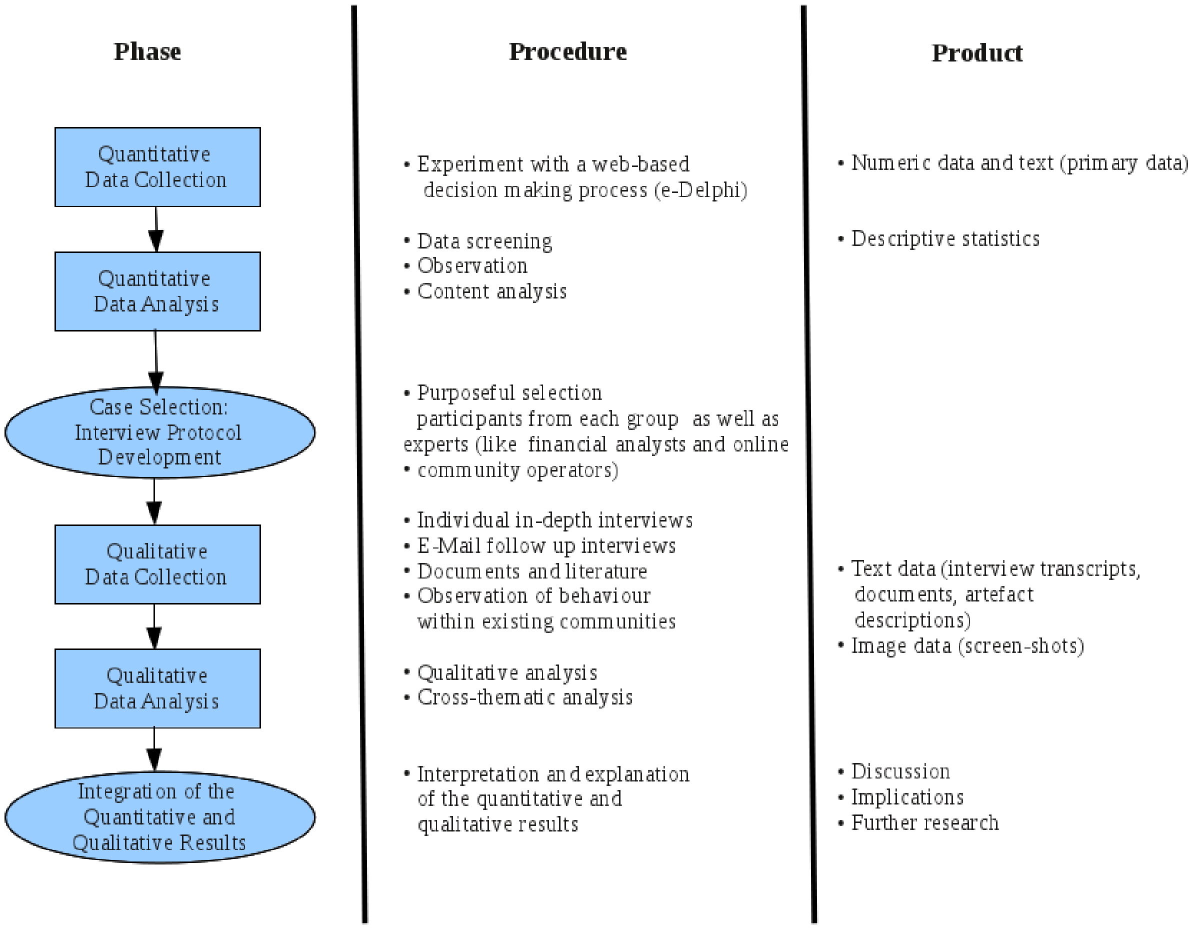


Figure 11. Methods and procedures

The Internet allowed fast and flexible access for participants of the survey. Figure 12 shows a mock-up of the survey form for the web-based Delphi method. The participants were requested to fill in a form like this for each of the different shares in the survey.

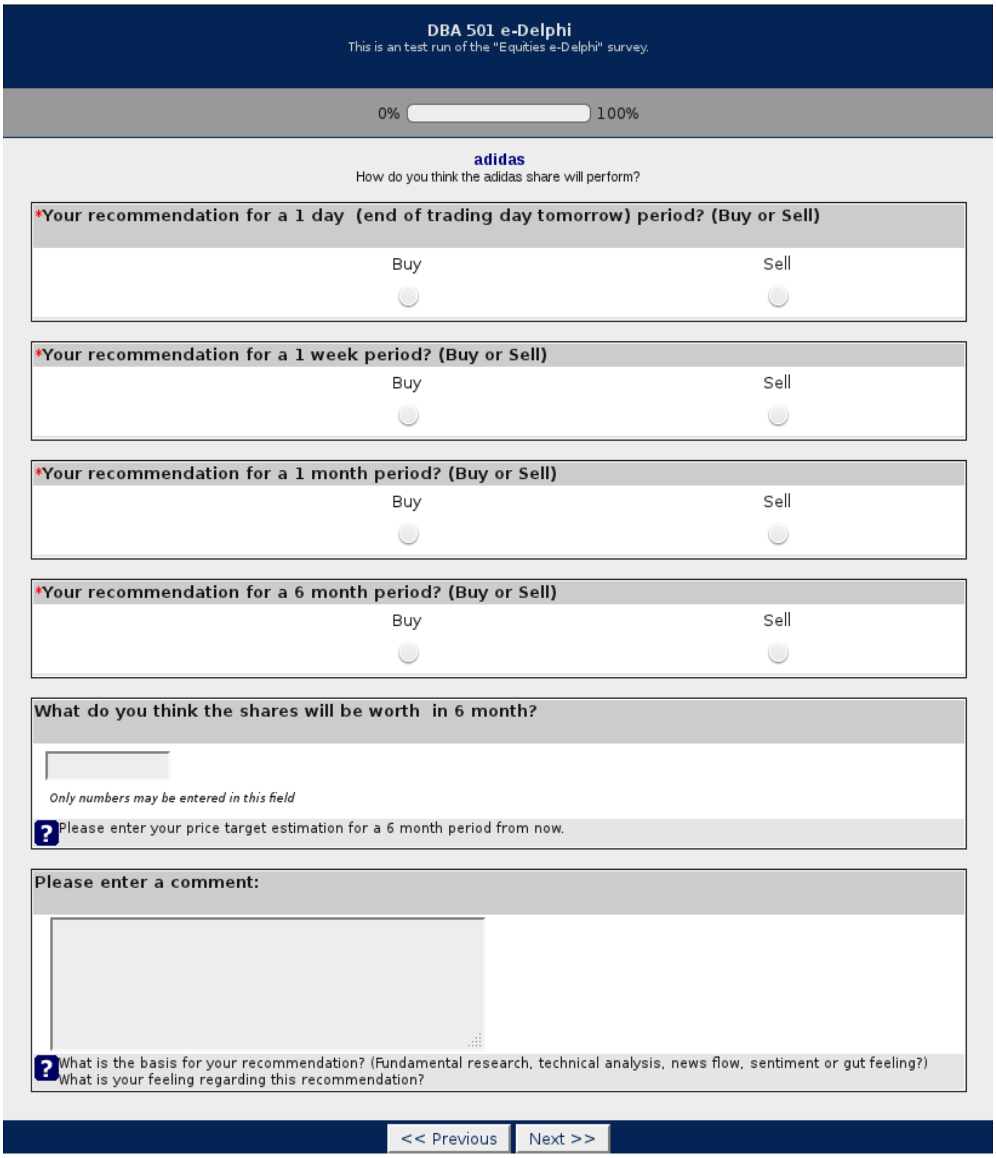


Figure 12. e-Delphi survey (first mockup)

3.8 Analysis Tools and Techniques

Generally the analysis of data generated in mixed-methods studies, like numbers and text data, involves the utilization of quantitative and qualitative analysis techniques (Combs & Onwuegbuzie, 2010). The use of quantitative and qualitative analysis methods aims in particular to allow a more holistic view (Creswell, 2009; Thomas, 2009) in the evaluation of the differences between individuals, the experimental groups, existing communities, the forecasts of financial analysts and actual market performance.

Quantitative analysis: factor analysis, statistical tests of significance, and time series analysis (Backhaus, Erichson, Plinke, & Weiber, 2016; Swift & Piff, 2005), aided by Excel, mySQL, SPSS and Preachers' Calculation for the chi-square test (Preacher, 2001). Preachers Calculation for the chi-square test is a website based interactive calculation tool for chi-square tests of goodness of fit and independence. These tests are primarily used to detect group differences using frequency data.

The usage of SPSS is based on the approach of Janssen & Laatz (2007), Backhaus et al. (2006; 2016) as well as the comments and help functions implemented in SPSS. One of the results of the statistical test methods applied is in each case a calculated p-value. The smaller the p-value, the greater the likelihood that a postulated difference between the samples actually exists. In this analysis the process was based on a threshold of $p \leq 0.05$ for a statistically significant result, i.e. that the established difference between groups or the relationship between two variables is not due to chance. According to Bortz (2005) this is the usual level of significance $p=0.05$ used (designation: “significant”). Test variables that provide a p-value between 0.05 and 0.1, are called “weakly significant”. A cluster analysis was also conducted with an SPSS, but the results were not conclusive and therefore not utilizable for the further analysis.

Qualitative analysis: narrative analysis “Considering the potential of stories to give meaning to individuals' lives, and treating data as stories, enabling researchers to take account of research participants' own evaluations” (Combs & Onwuegbuzie, 2010, p. 410). An analysis based on adapted disclosure/conversation analysis (DA/CA) techniques to identify recurring themes from the interviews (Bryman & Bell, 2003). According to Finlay (2002) one key to the validity of a mixed-methods study is reflexivity, because there is learning during the inquiry that possibly influences the process and outcomes. Reflection techniques have for many years been an established key to creating professional knowledge (Schön, 1983). Balton (2010) state that “reflection is a state of mind, an ongoing constituent of practice, not a technique, or curriculum element" (p. 3). This ongoing reflection during all stages of the sequential study targets the objective of creating and enhancing methods appropriate to developing an explanatory schema and starting theory building. The qualitative analysis was aided by MAXQDA, a software tool for qualitative and mixed-methods data analysis.

4 Pilot Experiment

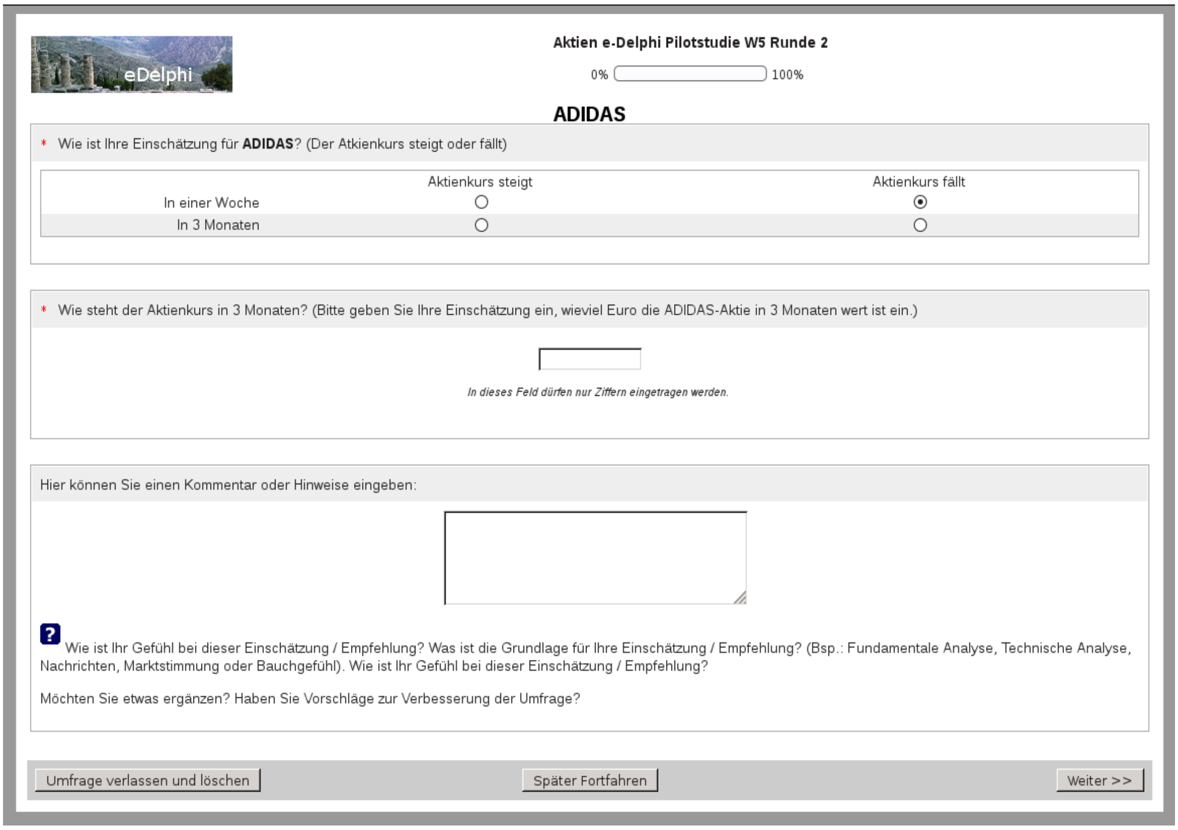
A pilot test of the operation of the online process for the proposed research was conducted using a small sample. The purpose of the study following the pilot is to gain an understanding of the group decision-making process of Internet communities, focusing on stock trading based on predicting share prices.

4.1 Pilot Stage Experiment Design

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 could provide different responses. 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 estimate of the movement (up or down) over a one-week and three-month period of every share as well as enter a stock price prediction for a three-month period (see Figure 13).

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. It was possible to handle the e-Delphi survey, given the set-up, software (Limesurvey) and Internet infrastructure chosen. The feedback from most participants was that the set-up was easy to use and the questions were easy to understand.


eDelphi

Aktien e-Delphi Pilotstudie W5 Runde 2

0% 100%

ADIDAS

* Wie ist Ihre Einschätzung für **ADIDAS**? (Der Atkienkurs steigt oder fällt)

	Aktienkurs steigt	Aktienkurs fällt
In einer Woche	○	◉
In 3 Monaten	○	○

* Wie steht der Aktienkurs in 3 Monaten? (Bitte geben Sie Ihre Einschätzung ein, wieviel Euro die ADIDAS-Aktie in 3 Monaten wert ist ein.)

In dieses Feld dürfen nur Ziffern eingetragen werden.

Hier können Sie einen Kommentar oder Hinweise eingeben:

? Wie ist Ihr Gefühl bei dieser Einschätzung / Empfehlung? Was ist die Grundlage für Ihre Einschätzung / Empfehlung? (Bsp.: Fundamentale Analyse, Technische Analyse, Nachrichten, Marktstimmung oder Bauchgefühl). Wie ist Ihr Gefühl bei dieser Einschätzung / Empfehlung?

Möchten Sie etwas ergänzen? Haben Sie Vorschläge zur Verbesserung der Umfrage?

Umfrage verlassen und löschen | Später Fortfahren | Weiter >>

Figure 13: Online Survey (Sample Screenshot of the Pilot Survey)

4.2 Findings, Value and Knowledge Contribution of the Pilot Run

The pilot run of the proposed experiment already provided a few indications that for an online group to make–in certain situations and with careful group design–predictions that are superior to predictions by experts might be possible. In particular, the pilot run helped identify the basic proceedings of the individuals' decision-making approaches. These preliminary results were the basis for the later survey design and allowed us to create clusters of different decision-making types. The results also indicated that there is some potential to improve the survey design, and adjust the structure and process slightly. In general, the pilot experiment demonstrated the feasibility of the experiment and showed that the tools and set-up are capable of conducting the proposed experiment.

The pilot experiment was aimed to gain a deeper understanding for the planned research later on. 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 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 3-month results generally confirmed the results from examination of the 1-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 which contributed to 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 often come up, and old ideas need to be redefined accordingly. One interesting idea as follow-up might be to test a group with a stronger feedback loop, such as 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.

4.2.1 Key Learning from the pilot experiment

The pilot experiment in generally demonstrated the feasibility of the research project and its suitability to address the research questions with the tested research design. The pilot experiment provided also some indication of how the research design might be slightly improved. The key learning from the pilot experiments can be categorized as follows:

- Adjust group design and feedback loop
- Assessment of the participants
- Enhancements of the online questionnaire

4.2.2 Adjust 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 5 weeks of the pilot.

Another finding of the pilot was that people did not change their predictions very often after receiving the group feedback with the e-Delphi method. The literature suggests more changes and a stronger convergence of the group decision (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 accordingly be to implement a stronger feedback loop for one group. This stronger feedback loop was facilitated by an audio conference (with Skype) between e-Delphi round one and two. A second control group was 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 might be possible to determine the effect of the feedback on the group's decision-making more clearly.

4.2.3 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 done for all participants in 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 the decision-making type might not be very easy for the participant to answer. An approach to addressing this question was developed by Cornelia Betsch (2004; Schunk & Betsch, 2006; Traufetter, 2009). She created and thoroughly tested a questionnaire to determine people's preference for intuition and/or deliberation. 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.

4.2.4 Enhancements of the online questionnaire.

The analysis of the procedure and the results of the pilot experiment also provided some suggestions for improving the online questionnaire. The questions about share movement (up or down) turned out to be useful and easy to understand, but not many participants provided information about their decision-making process in the free text-field on the online form. Nevertheless, the interviews of participants during the pilot run indicated some clusters of different types and sources for the decision-making process (see section 5.1 Main Stage Experiment Design and Quantitative Data Analysis on page 90). In order to simplify the answer options and to get more information these types were provided as a tick-box field for each of the participants' share estimate group, so that they might be more likely to provide more information about the background to their decision-making at the very moment they actually put their prediction into the online form. One participant in the pilot study did not feel comfortable with prediction of an actual price target for the 3 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 was changed into an optional question, in case anyone still feels uncomfortable with answering this question. Another change is to introduce a question about their level of confidence in their predictions (from not at all to absolutely sure, 1-5). Even though it might be interesting to include a couple 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 preferably remain simple to answer and to understand.

4.2.5 Participant interviews.

All participants in the pilot were interviewed. The questions 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 able to give estimates or at least enter a guess as to whether the stock price was going up or down. One participant felt uncomfortable about 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, 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 not as mandatory fields in the online survey, rather to leave it to the participants to enter a concrete price target with their prediction if they feel able and comfortable.

The interviews of the pilot experiment participants indicated different bases for the individual decisions. Here is an overview and a summary of the different answers, 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 did group in 9 clusters of different decision-making influences.

Table 8. *Clusters of different decision-making bases/influences*

Company	Products, brand, 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 1st 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	Index development, price-movement, momentum etc.

These clusters need to be transferred into easy to understand options for the lay participants in the main experiment. The participants are supposed to tick a box on the online survey for each weekly prediction for a company or add a comment if they used something not mentioned there.

5 Data and Analysis

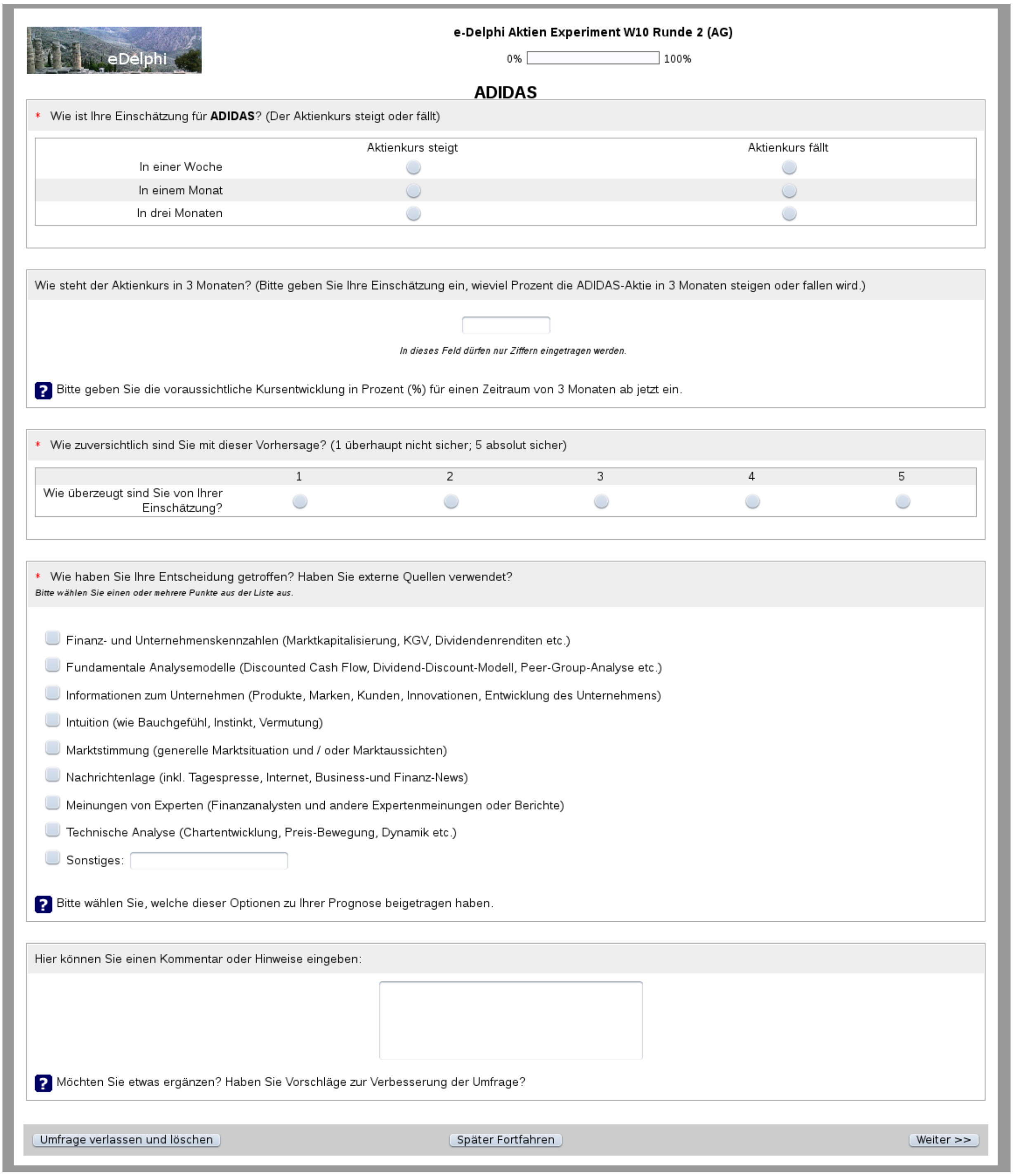
Learning from the pilot run led to an improved design of the main experiment. The main experiment was conducted using a bigger sample, a longer period, more shares and a wider range of different group designs. Additionally, the design of the questions (for the online survey as well as the semi-structured interviews) was slightly adjusted according to the suggestions and experiences from the pilot run.

5.1 Main Stage Experiment Design and Quantitative Data Analysis

Quantitative data analysis was conducted in a sequential approach. The first step was a univariate analysis. And in the second step a multi-criteria analysis and data reduction techniques were applied (with SPSS and Excel). Both approaches aim to inform an understanding of factors that influence the decision-making process and forecast quality.

The design and approach of the main experiment was in principle similar to the design of the pilot run (Endress, 2015). There were just a few changes in terms of an enhanced online questionnaire, more interview questions and the fact that they were asked to enter the target price not as the total amount in Euros but as a change in percent (see Figure 14).

There are numerous influencing factors which might impact investment decisions. These factors are—a side from measured variables and personal characteristics of the participants—variables such as risk aversion (Fellner-Röhling & Maciejovsky, 2007; Kahneman & Tversky, 1979; Keller & Siegrist, 2006), trading activity or sensation seeking (Grinblatt & Keloharju, 2009) which are the subject of discussion in behavioural finance literature. While it might be interesting to include these (and many other variables) in the experiment design, the measurement and discussion of all possible variables would limit the practicability of the research project. The proposed research design does not include trading activity or the calculation of (purely hypothetical) gains and losses. Eventually, it appears more adequate to assess the effects of these variables in context with investment decisions (including gain and loss calculations). This might be an interesting topic for further research. However, the experiment included an individual assessment of all participants including PID-score analysis (Betsch, 2004), personality traits (such as age, education etc.) and in-depth qualitative interviews.


eDelphi

e-Delphi Aktien Experiment W10 Runde 2 (AG)

0% 100%

ADIDAS

* Wie ist Ihre Einschätzung für **ADIDAS**? (Der Aktienkurs steigt oder fällt)

	Aktienkurs steigt	Aktienkurs fällt
In einer Woche		
In einem Monat		
In drei Monaten		

Wie steht der Aktienkurs in 3 Monaten? (Bitte geben Sie Ihre Einschätzung ein, wieviel Prozent die ADIDAS-Aktie in 3 Monaten steigen oder fallen wird.)

In dieses Feld dürfen nur Ziffern eingetragen werden.

? Bitte geben Sie die voraussichtliche Kursentwicklung in Prozent (%) für einen Zeitraum von 3 Monaten ab jetzt ein.

* Wie zuversichtlich sind Sie mit dieser Vorhersage? (1 überhaupt nicht sicher; 5 absolut sicher)

	1	2	3	4	5
Wie überzeugt sind Sie von Ihrer Einschätzung?					

* Wie haben Sie Ihre Entscheidung getroffen? Haben Sie externe Quellen verwendet?
Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

- Finanz- und Unternehmenskennzahlen (Marktkapitalisierung, KGV, Dividendenrenditen etc.)
- Fundamentale Analysemodelle (Discounted Cash Flow, Dividend-Discount-Modell, Peer-Group-Analyse etc.)
- Informationen zum Unternehmen (Produkte, Marken, Kunden, Innovationen, Entwicklung des Unternehmens)
- Intuition (wie Bauchgefühl, Instinkt, Vermutung)
- Marktstimmung (generelle Marktsituation und / oder Marktaussichten)
- Nachrichtenlage (inkl. Tagespresse, Internet, Business-und Finanz-News)
- Meinungen von Experten (Finanzanalysten und andere Expertenmeinungen oder Berichte)
- Technische Analyse (Chartentwicklung, Preis-Bewegung, Dynamik etc.)
- Sonstiges:

? Bitte wählen Sie, welche dieser Optionen zu Ihrer Prognose beigetragen haben.

Hier können Sie einen Kommentar oder Hinweise eingeben:

? Möchten Sie etwas ergänzen? Haben Sie Vorschläge zur Verbesserung der Umfrage?

Umfrage verlassen und löschen | Später Fortfahren | Weiter >>

Figure 14: Online Survey (Sample Screenshot of the Main Survey)

The main run was performed with 59 participants in three groups of lay people (21 participants, 21 participants, and 7 participants) and two groups with professionals, financial analysts, to benchmark the group over ten e-Delphi cycles (ten weeks with a two weeks break, i.e. the main data collection of the experiment was conducted during a 12-week period). The groups were as follows:

- Analyst Group (AG) with a group size of 5 participants
- e-Delphi-Group (EDG) with a group size of 21 participants
- Interactive Group (IG) with a group size of 7 participants
- Non-Feedback Group (NFG) with a group size of 21 participants
- Professional Investors Group (PG) with a group size of 5 participants

Additionally, the Single Expert/ Financial Analyst estimations were analysed as individual expert opinion within the narrow field of expertise in terms of active professional coverage of the respective company. Like the pilot run the main field experiment was conducted following an e-Delphi 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 within the groups to some groups (EDG, IG, PG) and, as control groups, two groups did not get any feedback from their group members (AG, NFG). In a second round, participants could provide different responses. The shares were selected from five different companies in four different sectors: consumer goods (Adidas, Bloomberg Symbol: ADS GY Equity), construction material (HeidelbergCement, Bloomberg Symbol: HEI GY Equity), utilities (RWE, Bloomberg Symbol: RWE GY Equity), industrial technology (Siemens, Bloomberg Symbol: SIE GY Equity), and industry (ThyssenKrupp, Bloomberg Symbol: TKA GY Equity). In all, the main experiment was set up to gather up to 17700 individual judgements about equity predictions (i.e. 5900 individual judgements about equity predictions for each period).

There was a quite bullish market condition in the relevant period of the main run. The DAX index went up about 17% during the examination period. Nevertheless, the different stocks had different price movements during the examination period (see Figure 15); while some stocks mostly went up (Adidas +36.56%, HeidelbergCement +43.04%), others went down (RWE -12.4%, ThyssenKrupp -8.67%) and one share showed a sidewards tendency and no clear direction (Siemens +5.18%). Each participant in the experiment was asked to provide an estimate of the movement of every share (up or down) over a one-week, a one-month and three-month period as well as enter a stock price change prediction in percent for a three-month period (see Figure 14). Group results with an undecided voting result, i.e. same number of votes for “up” and “down” were excluded. In some rounds, the groups came up with no recommendation (meaning that exactly 50% of the participants voted up and 50% voted down or the single expert vote was missing), and these undecided rounds have been excluded from the analysis. Missing votes from single experts were also excluded.

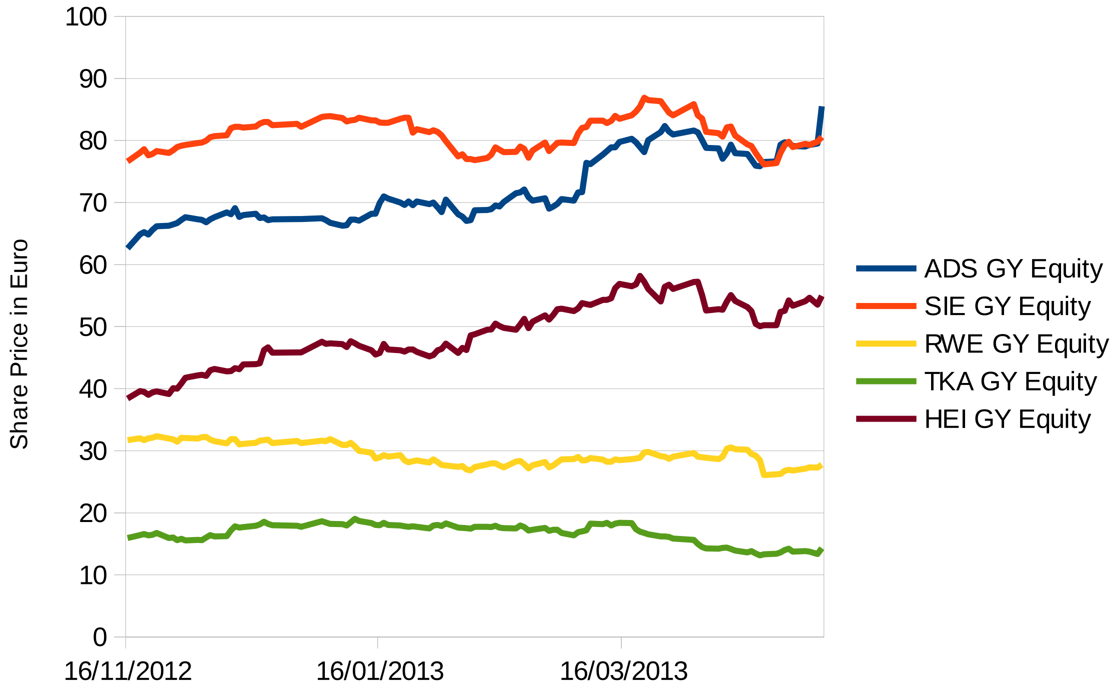


Figure 15: Share Price Development During the Main Experiment

5.1.1 Description of participants.

The following remarks represent the individuals surveyed in a summarized report. 60 participants agreed initially to participate in the experiment, but 59 people actually participated actively and provided valid answers.

19 participants did not have a university degree, while 40 did have a university degree.

10 participants are from the age group "up to 30 years", 25 participants "up to 40 years", 17 "up to 50 years" and seven "over 50 years". Most participants (N=17) categorized themselves as "rather rational", while almost the same number of participants categorized themselves as "emotional" (see also Table 9). There are no significant differences in the frequency of self-assessments in the different age groups (Pearson chi-square=5.742; DF=9; *p-value*=0.765).

Table 9. *Crosstab Self-assessment *Age Group*

		Age Group				Sum
		up to 30 Y.	up to 40 Y.	up to 50 Y.	over 50 Y.	
Self-Assessment	Emotional	3	7	4	2	16
	Rather emotional	2	5	4	0	11
	Rather rational	4	5	6	2	17
	Rational	1	8	2	2	13
Sum		10	25	16	6	57

People with university degrees were most common in the age group to 40 years (N = 21), whereas there were significantly fewer (per 3 respondents) in the age groups up to 30 years and over 50 years. In the age group up to 30 years, most respondents were without university degrees (N = 7). Significant differences in the incidence of university degree by age group are clearly detectable (Pearson chi-square=12.13; DF=3; *p-value*=0.007).

Table 10. *Crosstab University Degree *Age Group*

		Age Group				Sum
		up to 30 Y.	up to 40 Y.	up to 50 Y.	over 50 Y.	
University Degree	no	7	4	4	4	19
	yes	3	21	13	3	40
Sum		10	25	17	7	59

Participants with a university degree reported on average significantly higher values on the variable “Skill Self Estimation” than people without a university degree (see also Figure 16).

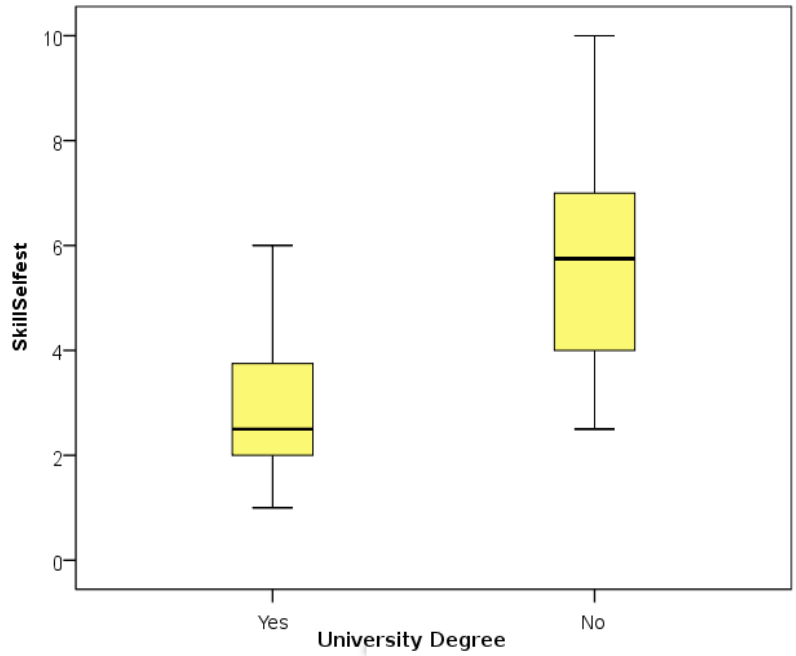


Figure 16. Skill Self-Assessment and University Degree (Box Plot)

Also, for the variable “Skill Self-Estimation” the values in the four age groups are considerably different. It is remarkable that young people and people above 50 assess their own skills lower than people in the age groups “up to 40” and “up to 50” (see Figure 17).

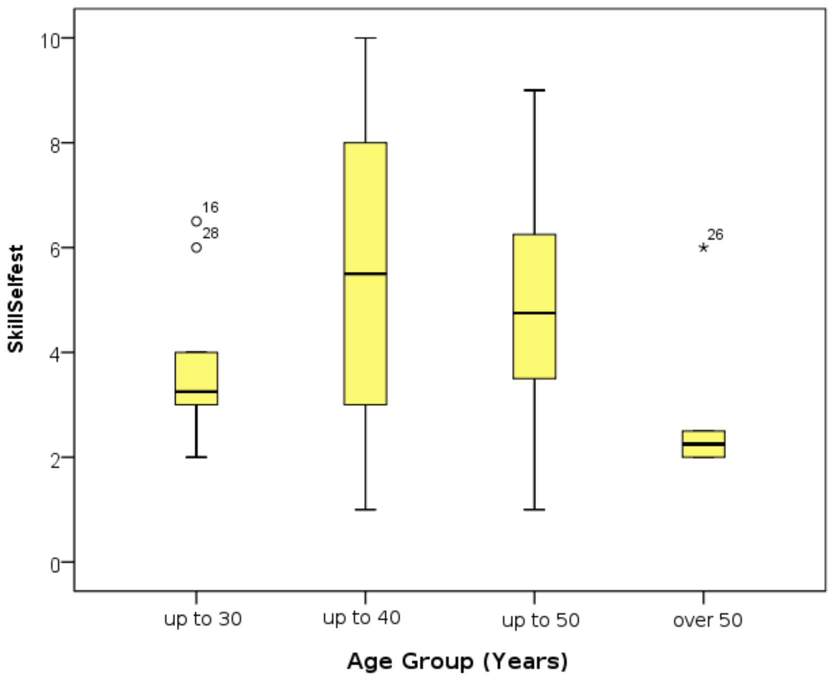


Figure 17. Skill Self-Assessment and Age Groups (Box Plot)

The more the respondents assess themselves as rational in terms of their decision-making, the higher their self-assessment of their skill in terms of knowledge about the stock market, represented by the value of the variable Skill Self-Estimation.

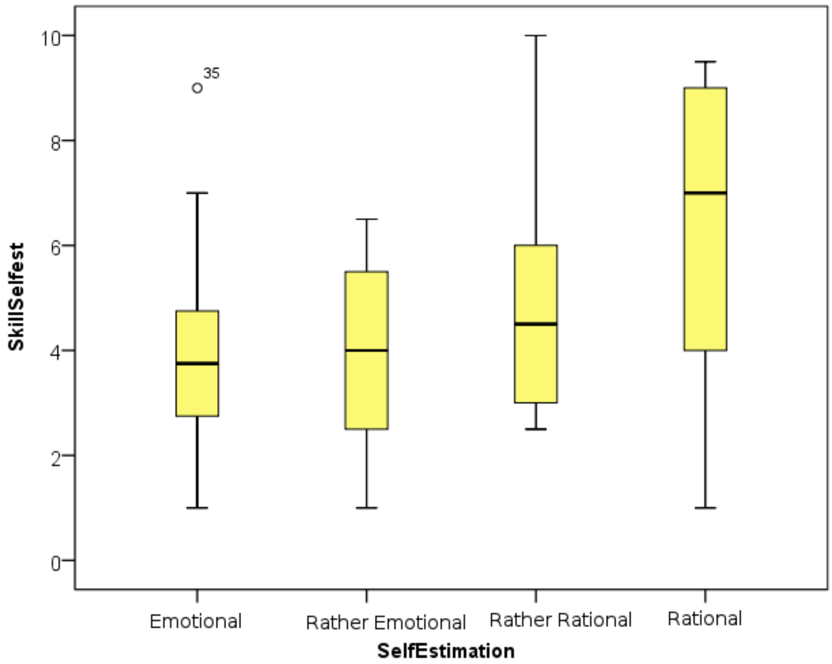


Figure 18. Self-Assessment of Skill and Rationality/Emotionality (Box Plot)

The absolute performance of the aggregated predictions from lay groups (52% correct predictions, see Table 11) was slightly above the value of 50% correct predictions, which would be the expected value with a purely random distribution (correct/wrong in the ratio 1:1). However, this outperformance is not significant (Chi-square 0.688; *p-value*=0.406).

The relative performance on an aggregated level from the lay groups compared with the expert recommendation were significantly different. The predictions by financial analysts were significantly better than those by lay groups (Chi-square 10.55; *p-value*=0.001). However, there were also considerable differences in predictive accuracy for the three different prediction periods.

Table 11. *Aggregated Predictive Accuracy from the Main Experiment*

	Sum (All Lay Groups)	Expert
Correct	407	54
Wrong	403	36
Excluded	60	10
Correct (%)	52.0%	63.3%
Wrong (%)	48.0%	36.7%

5.2 One-Week Predictions Main Stage

The main run of the group decision-making experiment confirmed the finding from the pilot that a mixed-method approach (Creswell, 2009; Johnson & Onwuegbuzie, 2004; Tashakkori, 2010) works in this context. The feedback from most participants confirmed the findings from the pilot run that the set-up was easy to use and the questions were easy to understand. The short term estimates (for one week) did not generally confirm that groups of lay people are better at predicting stock price movements than the experts (see Table 12). From 100 predictions (m=100), the e-Delphi-Group (EDG) had just 42 (43.8%) correct predictions, the financial analyst group (AG) had 48 (59.3%) correct predictions, the interactive group 41 (46.1%) correct predictions, the non-feedback group 57 (60.0%) correct predictions, and the single expert had 54 (60.0%) correct predictions.

Table 12. *Aggregated 1-Week Main Run Predictions*

	AG	EDG	IG	NFG	PG	Expert
Correct	48	42	41	57	30	54
Wrong	33	54	48	38	52	36
Excluded	19	4	11	5	18	10
Correct (%)	59.3%	43.8%	46.1%	60.0%	36.6%	60.0%
Wrong (%)	40.7%	56.3%	53.9%	40.0%	63.4%	40.0%

Generally, it can be noted that lay groups (EDG, IG, and NFG) did not perform per se better than the professionals (AG, PG, and individual experts), but the groups without feedback loop performed better, with 59.7% correct predictions overall (see Table 13), compared to the groups with feedback (EDG, IG, PG). This finding supports the idea that collective intelligence works best with diverse and independent group members (Page, 2008b).

Table 13. *Group Results with and without Feedback Loop for 1 Week Main Run Predictions*

		Groups without feedback	Groups with feedback loop
1 Week	Correct	105	113
	Wrong	71	154
	Excluded	24	33
	Correct (%)	59.7%	42.3%
	Wrong (%)	40.3%	57.7%

Not only the aggregated predictions of the group, but also the underlying individual decisions reveal significant differences between the groups (for all 3 periods) in terms of predictive accuracy. For the one week period chi-square 17.535, DF=4, *p-value*=0.002. The following crosstab provides an overview of the one week predictions (see Table 14).

Table 14. *Crosstab 1 Week Main Run Predictions*

			Group					
			AG	EDG	IG	NFG	PG	Sum
1 Week Predictions	wrong	Frequency	203	885	270	727	204	2289
		Expected Frequency	225.2	848.4	260.3	770.8	184.2	2289.0
		% in Group	45.1%	52.2%	51.9%	47.2%	55.4%	50.1%
	correct	Frequency	247	810	250	813	164	2284
		Expected Frequency	224.8	846.6	259.7	769.2	183.8	2284.0
		% in Group	54.9%	47.8%	48.1%	52.8%	44.6%	49.9%
Sum		Frequency	450	1695	520	1540	368	4573
		Expected Frequency	450.0	1695.0	520.0	1540.0	368.0	4573.0
		% in Group	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Considering all the participants' predictions individually for the one week prediction period (*N*=4573), the number of correct forecasts was about 49.9% and of wrong predictions about 50.1%. The AG group had the highest proportion of correct predictions (54.9%), followed by the NFG (52.8%). The remaining groups had less than 50% correct predictions.

5.3 One-Month Predictions Main Stage

The main run of the group decision experiment also included one-month predictions. The one-month prediction results did not confirm the idea that lay groups are better at predicting stock price movements than the experts (see Table 15). Of 100 predictions (m=100), the e-Delphi-Group (EDG) had just 42 (44.2%) correct predictions, the financial analyst group (AG) had 45 (51.1%) correct predictions, the interactive group 35 (41.2%) correct predictions, the non-feedback group 49 (52.1%) correct predictions, and the single expert had 57 (63.3%) correct predictions.

Table 15. *Aggregated One-Month Main Run Predictions*

	AG	EDG	IG	NFG	PG	Expert
Correct	45	42	35	49	56	57
Wrong	43	53	50	45	27	33
Excluded	12	5	15	6	17	10
Correct (%)	51.1%	44.2%	41.2%	52.1%	67.5%	63.3%
Wrong (%)	48.9%	55.8%	58.8%	47.9%	32.5%	36.7%

Generally, it can be noted that lay groups (EDG, IG, and NFG) did not perform per se better than the professionals (AG, PG, and individual experts), but the groups without feedback loop performed slightly better, with 51.6% correct predictions overall (see Table 16), compared to the groups with feedback (EDG, IG, PG).

Table 16. *Group Results with and without Feedback Loop for One Month Main Run Predictions*

		Groups without feedback	Groups with feedback loop
1 Month	Correct	94	133
	Wrong	88	130
	Excluded	18	37
	Correct (%)	51.6%	50.6%
	Wrong (%)	48.4%	49.4%

For the one-month predictions the underlying individual decisions also reveal significant differences in predictive accuracy between the groups: for the one-month period chi-square 18.794, DF=4, *p-value*=0.001. The following crosstab provides an overview of the one-month predictions (see Table 17).

Table 17. *Crosstab One-Month Main Run Predictions*

			Group					Sum
			AG	EDG	IG	NFG	PG	
1 Month Predictions	wrong	Frequency	204	901	275	749	162	2291
		Expected Frequency	225.4	849.2	260.5	771.5	184.4	2291.0
		% in Group	45.3%	53.2%	52.9%	48.6%	44.0%	50.1%
	correct	Frequency	246	794	245	791	206	2282
		Expected Frequency	224.6	845.8	259.5	768.5	183.6	2282.0
		% in Group	54.7%	46.8%	47.1%	51.4%	56.0%	49.9%
Sum		Frequency	450	1695	520	1540	368	4573
		Expected Frequency	450.0	1695.0	520.0	1540.0	368.0	4573.0
		% in Group	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

For the one-month prediction period all the participants predictions were considered individually (N=4573), the number of correct forecasts being about 55.6% and the wrong predictions about 44.4%. In the one-month period the professionals led the groups. The PG group had the highest proportion of correct predictions (56.0%), followed by the AG (54.7%) and NFG (51.4%). The remaining groups had less than 50% correct predictions.

5.4 Three-Month Predictions Main Stage

The 3-month prediction results were also (at least partially) contrary to the assumption that groups of lay people are better in predicting stock price movements than the experts (see Table 18). While from 100 predictions (m=100), the e-Delphi-Group (EDG) had 70 (72.2%) correct predictions, the financial analyst group (AG) had only 45 (47.7%) correct predictions, however, the interactive group just 42 (45.2%) correct predictions, the non-feedback group 59 (61.1%) correct predictions, and the single expert had 60 (66.7%) correct predictions. The best performance in the main run was from the financial professionals (PG) with a frequency of 62 (73.8%) correct predictions.

Table 18. *Aggregated 3-Month Main Run Predictions*

	AG	EDG	IG	NFG	PG	Expert
Correct	41	70	42	59	62	60
Wrong	46	27	51	37	22	30
Excluded	13	3	7	4	16	10
Correct (%)	47.1%	72.2%	45.2%	61.5%	73.8%	66.7%
Wrong (%)	52.9%	27.8%	54.8%	38.5%	26.2%	33.3%

Generally, it should be noted that lay groups (EDG, IG, and NFG) did not perform per se better than the professionals (AG, PG, and individual experts), but, in contrast to the 1-week and 1-month predictions, the groups with the feedback loop performed slightly better, with 63.5% correct predictions overall (see Table 19), compared to the groups without feedback (EDG, IG, PG).

Table 19. *Group Results with Feedback Loop and without Feedback for 1-Month Main Run Predictions*

		Groups without feedback	Groups with feedback loop
1 Month	Correct	100	174
	Wrong	83	100
	Excluded	17	26
	Correct (%)	54.6%	63.5%
	Wrong (%)	45.4%	36.5%

Also, for the three-month the underlying individual decisions reveal significant differences in predictive accuracy between the groups. The calculated significance value for the three-month period is chi-square 35.407, DF=4, *p-value*<0.0001. The following crosstab provides an overview of the three-month predictions (see Table 20).

Table 20. *Crosstab Three-Month Main Run Predictions*

			Group					Sum
			AG	EDG	IG	NFG	PG	
3-Month Predictions	wrong	Frequency	212	673	278	707	159	2029
		Expected Frequency	199.7	752.1	230.7	683.3	163.3	2029.0
		% in Group	47.1%	39.7%	53.5%	45.9%	43.2%	44.4%
	correct	Frequency	238	1022	242	833	209	2544
		Expected Frequency	250.3	942.9	289.3	856.7	204.7	2544.0
		% in Group	52.9%	60.3%	46.5%	54.1%	56.8%	55.6%
Sum		Frequency	450	1695	520	1540	368	4573
		Expected Frequency	450.0	1695.0	520.0	1540.0	368.0	4573.0
		% in Group	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

For the three-month prediction period, all the participants predictions' being considered individually, the number of correct forecasts was about 55.6% and the wrong predictions about 44.4%. The EDG group had the highest proportion of correct predictions (60.3%), followed by the PG (56.8%) and NFG (54.1%), the AG also still had more than 50% correct predictions. Only the IG had considerably less accuracy and only 46.5% correct predictions.

As a preliminary finding it can be noted that there are different groups of participants above the value of 50% correct predictions, which would be the expected value with a purely random distribution (correct/wrong in the ratio 1:1). These groups are:

- For 1-week periods: AG (54.9 %) and NFG (52.8%).
- For 1-month periods: PG (56.0 %). AG (54.7 %); NFG (51.4%).
- For 3-month periods: EDG (60.3 %); PG (56.8 %); NFG (54.1%); and AG (52.9 %).

It may be noteworthy that the groups NFG and AG are at all three time points higher than 50%, but not the PG group (only for 1 month and 3 months above 50%). However, the highest correct predictive accuracy was achieved by the EDG group with 60.3% for the 3-month predictions.

5.4.1 Comparison of predictions for different companies.

In the examination of the predictions for a specific company the group of lay people was better at predicting the stock price movement of the Adidas share than were the experts (see Table 21). Of 60 predictions (m = 60), the EDG group had 45 correct predictions (79%), 12 wrong predictions (21%), and in three 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 financial analyst group had 19 correct predictions (40%), and the single experts had 35 correct predictions (58%). An interesting observation might be that the financial analysts with coverage, i.e., in their narrow field of expertise had a considerably higher proportion of correct answers. The differences in the predictive accuracy of the groups are significant (Chi-square: 25.39, DF=5, *p-value*<0.001*)*.

Table 21. *Comparison of Aggregated Group Predictions for Adidas Share*

		AG	EDG	IG	NFG	PG	Expert
1 Week	Correct	7	10	9	11	5	11
	Wrong	7	8	9	7	10	9
	Excluded	6	2	2	2	5	0
1-Month	Correct	7	15	11	15	11	11
	Wrong	10	4	7	5	4	9
	Excluded	3	1	2	0	5	0
3-Month	Correct	4	20	12	20	16	13
	Wrong	11	0	6	0	2	7
	Excluded	5	0	2	0	2	0
1-Week	Correct	50%	56%	50%	61%	33%	55%
	Wrong	50%	44%	50%	39%	67%	45%
1-Month	Correct	41%	79%	61%	75%	73%	55%
	Wrong	59%	21%	39%	25%	27%	45%
3-Month	Correct	27%	100%	67%	100%	89%	65%
	Wrong	73%	0%	33%	0%	11%	35%

In predicting the stock price movement of HeidelbergCement the NFG was the best performing lay group (see Table 22). Of 60 possible predictions (m = 60), the NFG group had 31 correct predictions (55%). The financial analyst group had 15 correct predictions

(30%), and the individual experts had 49 correct predictions (85%). The differences in the predictive accuracy of the groups are significant (Chi-square: 43.157, DF=5, *p-value*<0.001*)*. For HeidelbergCement the performance of the individual experts was outstanding in the whole experiment, in particular the longer term predictions (1-month and 3-month predictions) were almost all correct.

Table 22. *Comparison of Aggregated Group Predictions for HeidelbergCement Share*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	9	7	7	12	6	12
	Wrong	7	13	10	7	9	7
	Excluded	4	0	3	1	5	1
1-Month	Correct	4	2	6	6	9	18
	Wrong	11	17	11	13	8	1
	Excluded	5	1	3	1	3	1
3-Month	Correct	2	18	9	13	18	19
	Wrong	16	2	10	5	0	0
	Excluded	2	0	1	2	2	1
1-Week	Correct	56%	35%	41%	63%	40%	63%
	Wrong	44%	65%	59%	37%	60%	37%
1-Month	Correct	27%	11%	35%	32%	53%	95%
	Wrong	73%	89%	65%	68%	47%	5%
3-Month	Correct	11%	90%	47%	72%	100%	100%
	Wrong	89%	10%	53%	28%	0%	0%

In predicting the stock price movement of the RWE share the IG was the best performing lay group (see Table 23). Of 60 possible predictions (m = 60), the IG group had 32 correct predictions (59%). The financial analyst group had 48 correct predictions (83%), and the single experts had 33 correct predictions (79%). The differences in the predictive accuracy of the groups are significant (Chi-square: 36.671, DF=5, *p-value*<0.001*)*. Both the individual analysts and the analyst group did very well for RWE. In contrast, the EDG had an overall predictive accuracy of only 34%.

Table 23. *Comparison of Aggregated Group Predictions for RWE Share*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	14	6	13	10	8	11
	Wrong	4	13	5	9	11	3
	Excluded	2	1	2	1	1	6
1-Month	Correct	17	4	12	10	7	11
	Wrong	3	14	5	8	8	3
	Excluded	0	2	3	2	5	6
3-Month	Correct	17	9	7	12	9	11
	Wrong	3	10	12	8	5	3
	Excluded	0	1	1	0	6	6
1-Week	Correct	78%	32%	72%	53%	42%	79%
	Wrong	22%	68%	28%	47%	58%	21%
1-Month	Correct	85%	22%	71%	56%	47%	79%
	Wrong	15%	78%	29%	44%	53%	21%
3-Month	Correct	85%	47%	37%	60%	64%	79%
	Wrong	15%	53%	63%	40%	36%	21%

In predicting the stock price movement of the Siemens share the EDG was again the best performing lay group (see Table 24). Of 60 possible predictions (*m* = 60), the EDG group had 32 correct predictions (56%). The financial analyst group had 24 correct predictions (51%), and the individual expert had 28 correct predictions (49%). The differences in the predictive accuracy of the groups are significant (Chi-square: 14.636, DF=5, *p-value*=0.012*)*. Most groups had around 50% correct predictions, only the IG (34% correct answers) was considerably below that level and the PG (71%) better than all the others.

Table 24. *Comparison of Aggregated Group Predictions for Siemens Share*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	9	11	7	14	5	12
	Wrong	6	8	11	5	10	7
	Excluded	5	1	2	1	5	1
1-Month	Correct	7	11	2	9	18	8
	Wrong	11	8	15	8	1	11
	Excluded	2	1	3	3	1	1
3-Month	Correct	8	10	9	4	13	8
	Wrong	6	9	9	14	4	11
	Excluded	6	1	2	2	3	1
1-Week	Correct	60%	58%	39%	74%	33%	63%

	Wrong	40%	42%	61%	26%	67%	37%
1-Month	Correct	39%	58%	12%	53%	95%	42%
	Wrong	61%	42%	88%	47%	5%	58%
3-Month	Correct	57%	53%	50%	22%	76%	42%
	Wrong	43%	47%	50%	78%	24%	58%

In predicting the stock price movement of the ThyssenKrupp share the EDG was again the best performing lay group (see Table 25). Of 60 possible predictions (*m* = 60), the EDG group had 31 correct predictions (53%). The financial analyst group had 29 correct predictions (52%), and the individual expert had 22 correct predictions (48%). The IG had the lowest predictive accuracy for the ThyssenKrupp share (only 14 correct predictions; 26%). However, the differences in the predictive accuracy of the groups are not significant (Chi-square: 10.237, DF=5, *p-value*=0.069). The overall predictive accuracy was lowest for the ThyssenKrupp share.

Table 25. *Comparison of Aggregated Group Predictions for ThyssenKrupp Share*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	9	8	5	10	6	8
	Wrong	9	12	13	10	12	10
	Excluded	2	0	2	0	2	2
1-Month	Correct	10	10	4	9	11	9
	Wrong	8	10	12	11	6	9
	Excluded	2	0	4	0	3	2
3-Month	Correct	10	13	5	10	6	9
	Wrong	10	6	14	10	11	9
	Excluded	0	1	1	0	3	2
1-Week	Correct	50%	40%	28%	50%	33%	44%
	Wrong	50%	60%	72%	50%	67%	56%
1-Month	Correct	56%	50%	25%	45%	65%	50%
	Wrong	44%	50%	75%	55%	35%	50%
3-Month	Correct	50%	68%	26%	50%	35%	50%
	Wrong	50%	32%	74%	50%	65%	50%

5.5 Performance of the Individual Participants

The next table shows the performance of the individual members of the groups and their self-estimated knowledge about the stock market (scale 1-10, from 1=no knowledge to 10=expert). The initial analyses of the individual results showed that 27 of 49 participants had a success rate of 50% correct predictions (see Table 26. Main Run Predictions of Lay-Participants) or higher.

Table 26. *Main Run Predictions of Lay-Participants*

	Group	**Predictive Accuracy (ALL)**	**Predictive Accuracy (1W)**	**Predictive Accuracy (1M)**	**Predictive Accuracy (3M)**	**Skill self-assessment**
Participant 503	NFG	65.0%	56.0%	68.0%	71.0%	6
Participant 511	NFG	63.2%	56.8%	62.1%	70.5%	4
Participant 516	NFG	62.1%	60.0%	65.3%	61.1%	7.5
Participant 604	EDG	60.4%	56.7%	50.0%	74.4%	4
Participant 603	EDG	60.0%	50.0%	65.0%	65.0%	
Participant 508	NFG	60.0%	56.4%	65.5%	58.2%	3
Participant 510	NFG	59.6%	50.7%	57.3%	70.7%	3
Participant 620	EDG	59.4%	46.7%	68.3%	63.3%	1
Participant 601	EDG	58.5%	49.2%	66.2%	60.0%	5.5
Participant 517	NFG	58.3%	52.0%	59.0%	64.0%	2.5
Participant 613	EDG	57.9%	56.3%	58.8%	58.8%	4.5
Participant 618	EDG	57.0%	53.3%	56.7%	61.1%	4.5
Participant 615	EDG	56.7%	45.0%	51.0%	74.0%	3.5
Participant 512	NFG	56.4%	56.0%	53.3%	60.0%	1
Participant 519	NFG	56.3%	54.0%	55.0%	60.0%	2.5
Participant 614	EDG	55.9%	52.2%	47.8%	67.8%	3
Participant 518	NFG	55.4%	52.5%	50.0%	63.8%	2.5
Participant 36	IG	53.3%	38.2%	60.0%	61.8%	7
Participant 621	EDG	53.3%	42.4%	47.1%	70.6%	2.5
Participant 514	NFG	52.7%	56.0%	49.0%	53.0%	2
Participant 502	NFG	52.4%	57.1%	51.4%	48.6%	4
Participant 606	EDG	51.7%	42.5%	48.8%	63.8%	6
Participant 607	EDG	51.7%	50.0%	39.0%	66.0%	2
Participant 501	NFG	51.3%	67.0%	49.0%	38.0%	6.5
Participant 38	IG	50.7%	50.0%	49.0%	53.0%	5.5
Participant 605	EDG	50.5%	60.0%	45.7%	45.7%	4
Participant 608	EDG	50.0%	41.0%	41.0%	68.0%	2.5
Participant 616	EDG	49.6%	44.4%	32.2%	72.2%	4
Participant 34	IG	49.6%	46.3%	47.5%	55.0%	2
Participant 4	IG	48.6%	42.9%	51.4%	51.4%	
Participant 602	EDG	48.2%	44.7%	45.9%	54.1%	6
Participant 15	IG	48.2%	48.9%	50.0%	45.6%	4
Participant 513	NFG	47.3%	56.0%	37.0%	49.0%	2.5
Participant 612	EDG	47.0%	48.4%	47.4%	45.3%	3
Participant 610	EDG	46.7%	38.9%	42.2%	58.9%	4
Participant 14	IG	46.7%	53.3%	45.0%	41.7%	6
Participant 611	EDG	46.0%	49.0%	41.0%	48.0%	3
Participant 505	NFG	45.8%	47.5%	45.0%	45.0%	5
Participant 619	EDG	45.3%	44.0%	44.0%	48.0%	2
Participant 504	NFG	45.2%	50.0%	34.3%	51.4%	1
Participant 506	NFG	45.0%	31.3%	51.3%	52.5%	6.5
Participant 609	EDG	44.2%	43.8%	30.0%	58.8%	5
Participant 509	NFG	43.1%	58.5%	41.5%	29.2%	7

Participant 617	EDG	42.0%	47.0%	42.0%	37.0%	6
Participant 507	NFG	41.0%	37.1%	42.9%	42.9%	4.5
Participant 520	NFG	40.0%	40.0%	40.0%	40.0%	5
Participant 2	IG	37.7%	51.0%	35.0%	27.0%	6
Participant 515	NFG	37.5%	30.0%	45.0%	37.5%	3
Participant 521	NFG	36.1%	46.3%	32.6%	29.5%	3

Most participants missed a few of the 20 e-Delphi rounds (=10 weeks x 2 rounds), but there were only 2 “drop outs” in terms of a participant leaving the panel during the main experiment without returning. All but the 2 drop outs were interviewed in parallel or shortly after the e-Delphi rounds. In the interviews, all participants were asked to give a self-assessment of their investment expertise on a scale from 1 to 10 (1 = no knowledge; 10 = expert). It might be hypothesized that there would be a high correlation between success rate and self-estimated skill (see also discussion of assumption “A11” in the analysis part, page 190). Table 27 shows the predictive accuracy of the individual predictions of the experts (financial analysts and other experts) for the main run estimates.

Table 27. *Main Run Predictions of Analysts and Financial Professional-Participants*

	Group	**Predictive Accuracy (ALL)**	**Predictive Accuracy (1W)**	**Predictive Accuracy (1M)**	**Predictive Accuracy (3M)**	**Skill self-assessment**
Participant 204	AG	69.8%	62.1%	71.6%	75.8%	10
Participant 101	PG	60.4%	47.8%	64.4%	68.9%	9
Participant 102	PG	58.5%	44.6%	61.5%	69.2%	3
Participant 201	AG	55.7%	54.3%	61.4%	51.4%	8
Participant 202	AG	50.7%	48.0%	47.0%	57.0%	9
Participant 205	AG	49.6%	53.3%	50.0%	45.6%	9
Participant 104	PG	49.2%	31.8%	55.7%	60.2%	9
Participant 103	PG	48.9%	46.7%	57.8%	42.2%	6
Participant 203	AG	45.3%	56.8%	45.3%	33.7%	9.5
Participant 105	PG	44.2%	53.8%	41.3%	37.5%	8

5.6 Changes from First to Second e-Delphi Round

The group's overall decisions did not change fundamentally from the first to the second e-Delphi round of the main experiment (see Tables 28 and 29). The overall accuracy of the groups in the 1st e-Delphi rounds was about 56% correct predictions. The accuracy of the groups with feedback loop in the 1st e-Delphi rounds was about 53% correct answers (see Table 28).

Table 28. *Main Run Predictions in e-Delphi Round 1*

		AG	EDG	IG	NFG	PG	Expert
1Week	Correct	21	22	20	29	14	29
	Wrong	18	27	23	18	24	15
	Excluded	11	1	7	3	12	6
1-Month	Correct	24	21	21	26	25	27
	Wrong	19	27	20	22	10	17
	Excluded	7	2	9	2	15	6
3-Month	Correct	20	34	19	30	26	29
	Wrong	23	14	25	17	11	15
	Excluded	7	2	6	3	13	6
1-Week	Correct	54%	45%	47%	62%	37%	66%
	Wrong	46%	55%	53%	38%	63%	34%
1-Month	Correct	56%	44%	51%	54%	71%	61%
	Wrong	44%	56%	49%	46%	29%	39%
3-Month	Correct	47%	71%	43%	64%	70%	66%
	Wrong	53%	29%	57%	36%	30%	34%

The overall predictive accuracy of the groups in the 2nd e-Delphi rounds was about 54% correct answers and the accuracy of the groups with feedback loop about 52% correct answers (see Table 29). For further discussion on group learning effects please refer also to section “Area of Discussion A2: An improvement in predictive accuracy results from feedback from an e-Delphi group.“ on page 157 which provides further details and analysis on possible improvements and differences between 1st and 2nd e-Delphi round.

Table 29. *Main Run Predictions in e-Delphi Round 2*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	27	20	21	28	16	25
	Wrong	15	27	25	20	28	21
	Excluded	8	3	4	2	6	4
1-Month	Correct	21	21	14	23	31	30
	Wrong	24	26	30	23	17	16
	Excluded	5	3	6	4	2	4
3-Month	Correct	21	36	23	29	36	31
	Wrong	23	13	26	20	11	15
	Excluded	6	1	1	1	3	4
1-Week	Correct	64%	43%	46%	58%	36%	54%
	Wrong	36%	57%	54%	42%	64%	46%
1-Month	Correct	47%	45%	32%	50%	65%	65%
	Wrong	53%	55%	68%	50%	35%	35%

3-Month	Correct	48%	73%	47%	59%	77%	67%
	Wrong	52%	27%	53%	41%	23%	33%

5.6.1 Group learning during the main run.

Table 30. *Predictions from Week 1-5 (1st Half)*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	19	14	20	30	11	23
	Wrong	18	33	30	19	24	20
	Excluded	13	3	0	1	15	7
1-Month	Correct	17	7	21	16	31	20
	Wrong	24	40	29	31	7	23
	Excluded	9	3	0	3	12	7
3-Month	Correct	11	33	22	20	32	20
	Wrong	28	15	28	27	4	23
	Excluded	11	2	0	3	14	7
1-Week	Correct	51%	30%	40%	61%	31%	53%
	Wrong	49%	70%	60%	39%	69%	47%
1-Month	Correct	41%	15%	42%	34%	82%	47%
	Wrong	59%	85%	58%	66%	18%	53%
3-Month	Correct	28%	69%	44%	43%	89%	47%
	Wrong	72%	31%	56%	57%	11%	53%

Table 31. *Predictions from Week 6-10 (2nd Half)*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	29	28	21	27	19	31
	Wrong	15	21	18	19	28	16
	Excluded	6	1	11	4	3	3
1-Month	Correct	28	35	14	33	25	37
	Wrong	19	13	21	14	20	10
	Excluded	3	2	15	3	5	3
3-Month	Correct	30	37	20	39	30	40
	Wrong	18	12	23	10	18	7
	Excluded	2	1	7	1	2	3
1-Week	Correct	66%	57%	54%	59%	40%	66%
	Wrong	34%	43%	46%	41%	60%	34%
1-Month	Correct	60%	73%	40%	70%	56%	79%
	Wrong	40%	27%	60%	30%	44%	21%
3-Month	Correct	63%	76%	47%	80%	63%	85%
	Wrong	38%	24%	53%	20%	38%	15%

5.6.2 Multivariate and factor analysis

The following section provides the results of a discriminant analysis to differentiate the correct or incorrect predictions depending on specific characteristics of the prediction (in terms of variables of the data set). Multivariate analysis allows one to examine several variables simultaneously and to describe and explain correlations (Backhaus et al., 2016). The discriminant analysis, a special form of multivariate analysis, allows the analysis of group differences with a huge number of variables (Backhaus et al., 2016). An analysis of the predictions to test the equality of group means (for groups with correct and incorrect predictions) indicates that some variables reveal significant differences between correct and wrong predictions, but there are also a number of variables with no significant differences. However, the explanatory power of the analysed variable assignment is limited (as it explains only a small percentage of the variability).

5.6.2.1 Discriminant analysis for one-week predictions.

An analysis of the one-week predictions to test the equality of group means (for groups with correct and incorrect predictions) indicates that some variables reveal significant differences between correct and wrong predictions (column “Signif.”), e.g. “commitment” (COMSQ001), “use of fundamental analysis” (SQ002) and “use of group results/feedback” (SQ003) etc. (with level of significance with p-values less than 0.05), but there are also a number of variables with no significant differences (see Table 32).

Table 32. *Tests of Equality of Group Means (1-Week Predictions)*

	Wilks-Lambda	F	df1	df2	Signif.
COMSQ001	.999	5.964	1	4487	.015
SQ001	1.000	.082	1	4487	.774
SQ002	.999	4.275	1	4487	.039
SQ003	.999	4.884	1	4487	.027
SQ004	1.000	.053	1	4487	.817
SQ005	1.000	1.563	1	4487	.211
SQ006	.999	4.978	1	4487	.026

SQ008	1.000	1.481	1	4487	.224
SQ007	.999	3.730	1	4487	.054
SQ009	1.000	1.682	1	4487	.195
Age Group	.998	6.744	1	4487	.009
PID-D	1.000	.407	1	4487	.523
PID-I	1.000	.324	1	4487	.569
Emo-Selfasses.	1.000	.169	1	4487	.681
Skill Selfasses.	1.000	.621	1	4487	.431
PA-ALL	.997	15.266	1	4487	.000
PA-1W	.979	98.276	1	4487	.000
PA-1M	.999	5.595	1	4487	.018
PA-3M	1.000	.795	1	4487	.373
Survey Share	1.000	.032	1	4487	.858
Group No	1.000	.262	1	4487	.609
Univ. Degree	1.000	.579	1	4487	.447

Using the tests with Wilks Lambda we examined whether the average discriminant scores of the two groups were different. As shown by the following two tables, approximately 15.1% of the variability of the discriminant scores is explained by the differences between the two groups (correct and wrong predictions). The statistical test of Wilk's Lambda is significant (*p-value*<0.0001), i.e. the variables can be used to differentiate the two groups statistically (Backhaus et al., 2016).

Table 33. *Eigenvalues and Wilks-Lambda (1-Week Predictions)*

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.023[a]	100.0	100,0	.151

a. First 1 canonical discriminant functions were used in the analysis.

Wilks-Lambda

Test of Function(s)	Wilks-Lambda	Chi-square	df	Sig.
1	.977	103.261	21	.000

The following table of standardized discriminant allows (significant variable in bold; variables that have not passed the tolerance test are not listed) comparisons between the variables, i.e., since the variables with a significant influence on the distinctness of the groups have large absolute value of the standardized discriminant. Accordingly, the variable “use of financial ratios” (SQ001), “use fundamental analysis” (SQ002) are relatively significant (values approximately at 0.15), the variables market sentiment (SQ006), group results (SQ003), age group and overall prediction accuracy (PAALL) are relatively unimportant (values below 0.1), and the variable predictive accuracy one week (PA1W) most important (highest value of the standardized discriminant), see the following table 34.

Table 34. *Standardized Canonical Discriminant Function Coefficients (1-Week Predictions)*

	Function
	1
COMSQ001	**-.152**
SQ001	-.032
SQ002	**.141**
SQ003	**-.075**
SQ004	-.033
SQ005	-.097
SQ006	**-.100**
SQ008	-.067
SQ007	.078
SQ009	.072
Age Group	**.045**
PID-D	-.007
PID-I	-.011
Emo-Selfasses.	-.043
Skill Selfasses.	.049
PAALL	**-.008**
PA1W	**.906**
PA1M	**-.084**
Survey Share	-.024
Group No	-.038
Univ. Degree	-.004

The mean of the discriminant for both groups is shown by the following table (see Table 35); the average value for wrong predictions is -0.153 and for correct predictions +0.153, i.e. a relatively small difference (possible values are from -1 to 1).

Table 35. *Functions at Group Centroids (1-Week Predictions)*

1-Week Predictions	Function 1
wrong	-.153
correct	.153

Unstandardised canonical discriminant functions evaluated at group means

The following table shows the results of the real and the predicted (using the discriminant function) assignment of groups (correct and wrong predictions). Only 56.1% of the original grouped cases were correctly classified. This is just a little above the random value (50%). This shows that the explanatory power of the existing variable assignment is rather limited (explains only 15% of the variability):

Table 36. *Classification Results (1-Week Predictions)*

		1-week predictions	Predicted Group Membership		Sum
			wrong	correct	
Original	Frequency	wrong	1205	1067	2272
		correct	928	1344	2272
	%	wrong	53.0	47.0	100,0
		correct	40.8	59.2	100.0

a. 56.1% of original grouped cases correctly classified.

5.6.2.2 Discriminant analysis for one-month predictions.

An analysis of the one-month predictions to test the equality of group means (for groups with correct and incorrect predictions) indicates that some variables reveal significant differences between correct and wrong predictions (column “Signif.”), e.g. commitment (COMSQ001), use of financial ratios (SQ001), use of expert opinions (SQ007), and technical analysis (SQ009) etc. (with level of significance with p-values less than 0.05), but there are also a number of variables with no significant differences.

Table 37. *Tests of Equality of Group Means (1-Month Predictions)*

	Wilks-Lambda	F	df1	df2	Signif.
COMSQ001	.999	6.542	1	4487	.011
SQ001	.999	4.456	1	4487	.035
SQ002	.999	2.245	1	4487	.134
SQ003	1.000	.236	1	4487	.627
SQ004	1.000	.828	1	4487	.363
SQ005	1.000	1.828	1	4487	.176
SQ006	.999	3.379	1	4487	.066
SQ008	1.000	1.315	1	4487	.252
SQ007	.998	8.831	1	4487	.003
SQ009	.996	18.454	1	4487	.000
Age Group	1.000	.062	1	4487	.804
PID-D	.998	7.833	1	4487	.005
PID-I	1.000	.176	1	4487	.675
Emo Self-Asses.	1.000	.928	1	4487	.336
Skill Self-Asses.	.998	10.648	1	4487	.001
PAALL	.973	126.546	1	4487	.000
PA1W	.998	11.145	1	4487	.001
PA1M	.961	184.320	1	4487	.000
PA3M	.991	40.728	1	4487	.000
SurveyShare1	1.000	.731	1	4487	.393
GroupNo	.999	3.299	1	4487	.069
University Degree	.996	18.437	1	4487	.000

Using the tests with Wilks Lambda we examined whether the average discriminant scores of the two groups were different. As shown by the following two tables, approximately 20.5% of the variability of the discriminant scores is explained by the differences between the two groups (correct and wrong predictions). The statistical test of Wilk's Lambda is significant (p-value <0.0001), i.e. the variables can be used to differentiate the two groups statistically (Backhaus et al., 2016).

Table 38. *Eigenvalues and Wilks-Lambda (1-Month Predictions)*

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.044[a]	100.0	100.0	.205

a. First 1 canonical discriminant functions were used in the analysis.

Wilks-Lambda

Test of Function(s)	Wilks-Lambda	Chi-square	df	Sig.
1	.958	191.438	21	.000

The following table of standardized discriminant allows (significant variable in bold; variables that have not passed the tolerance test are not listed) comparisons between the variables, i.e., since the variables with a significant influence on the distinctness of the groups have large absolute value of the standardized discriminant. Accordingly, the variables commitment (COMSQ001), “use of expert opinions” (SQ007) are relatively significant (values higher than 0.10), the variables “use of financial ratios” (SQ001), “use of technical analysis” (SQ009), “preference for deliberation” (PID-D), “self assessment of skill in the equity market” (Skill Selfasses.), “overall predictive accuracy” (PAALL) and “predictive accuracy one month” (PA1M) are relatively unimportant (values below 0.10), and the variable “predictive accuracy 1 month” (PA1M) is the most important (highest value of the standardized discriminant), see the following table 39.

Table 39. *Standardized Canonical Discriminant Function Coefficients (1-Month Predictions)*

	Function
	1
COMSQ001	**-.108**
SQ001	**-.050**
SQ002	.201
SQ003	.084
SQ004	-.061
SQ005	.054
SQ006	-.013
SQ008	.050
SQ007	**-.132**
SQ009	**.103**
Age Group	.042
PID-D	**.004**
PID-I	-.061
Emo. Self-Asses.	-.066
Skill Self-Asses.	**.024**
PAALL	**.061**
PA1W	**-.005**
PA1M	**.918**
Survey Share	.061
Group No	-.007
Univ. Degree	**.002**

The mean of the discriminant for both groups is shown by the following table (see Table 40): the average value for wrong predictions is -0.208 and for correct predictions +0.209, i.e. a relatively small difference (possible values are from -1 to 1).

Table 40. *Functions at Group Centroids (1-Month Predictions)*

1-Month Predictions	Function
	1
wrong	-.208
correct	.209

Unstandardised canonical discriminant functions evaluated at group means

The following table shows the results of the real and the predicted (using the discriminant function) assignment of groups (correct and wrong predictions). Only 58.5% of the original grouped cases were correctly classified. This is just a little above the random value (50%). This shows that explanatory power of the existing variable assignment is rather limited (explains only 20.5% of the variability):

Table 41. *Classification Results (1-Month Predictions)*

		1-week predictions	Predicted Group Membership		
			Wrong	correct	Sum
Original	Frequency	wrong	1463	811	2274
		correct	1073	1197	2270
	%	wrong	64.3	35.7	100.0
		correct	47.3	52.7	100.0

a. 58.5% of original grouped cases correctly classified.

5.6.2.3 Discriminant analysis for three-month predictions.

An analysis of the three-month predictions to test the equality of group means (for groups with correct and incorrect predictions) indicates that some variables reveal significant differences between correct and wrong predictions (column “Signif.”), e.g. “use of group results” (SQ003), “use of intuition” (SQ005) and “listen to market sentiment” (SQ006) etc. (with level of significance with p-values less than 0.05), but there are also a number of variables with no significant differences.

Table 42. *Tests of Equality of Group Means (3-Month Predictions)*

	Wilks-Lambda	F	df1	df2	Signif.
COMSQ001	1.000	.011	1	4487	.918
SQ001	1.000	.519	1	4487	.471
SQ002	1.000	2.139	1	4487	.144
SQ003	.998	10.465	1	4487	.001
SQ004	.999	2.688	1	4487	.101
SQ005	.997	11.862	1	4487	.001
SQ006	.997	12.501	1	4487	.000
SQ008	1.000	.001	1	4487	.977
SQ007	.998	8.073	1	4487	.005
SQ009	.999	3.718	1	4487	.054
Age Group	1.000	1.255	1	4487	.263
PID-D	.997	12.023	1	4487	.001
PID-I	.999	2.697	1	4487	.101
Emo. Self-Asses.	.997	14.131	1	4487	.000
Skill Self-Asses.	.998	10.723	1	4487	.001
PA-ALL	.958	194.524	1	4487	.000
PA-1W	1.000	2.158	1	4487	.142
PA-1M	.985	69.450	1	4487	.000
PA-3M	.933	320.817	1	4487	.000
Survey Share	.979	94.398	1	4487	.000
Group No	.998	6.787	1	4487	.009
Univ. Degree	1.000	.775	1	4487	.379

Using the tests with Wilks Lambda we examined whether the average discriminant scores of the two groups were different. As shown by the following two tables, approximately 29.8% of the variability of the discriminant scores is explained by the differences between the two groups (correct and wrong predictions). The statistical test of Wilk's Lambda is significant (p-value <0.0001), i.e. the variables can be used to differentiate the two groups statistically (Backhaus et al., 2016).

Table 43. *Eigenvalues and Wilks-Lambda (3-Month Predictions)*

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.043[a]	100.0	100,0	.204

a. First 1 canonical discriminant functions were used in the analysis.

Wilks-Lambda

Test of Function(s)	Wilks-Lambda	Chi-square	df	Sig.
1	.958	190.575	21	.000

The following table of standardized discriminant allows (significant variable in bold; variables that have not passed the tolerance test are not listed) comparisons between the variables, i.e., since the variables with a significant influence on the distinctness of the groups have large absolute value of the standardized discriminant. Accordingly, the variables “prediction accuracy one month” (PA1M), “company share” (Survey Share) are relatively important (values considerable higher than 0.40), the variables “use of group results” (SQ003), “use of intuition” (SQ005), “use of expert opinions” (SQ007), “preference for deliberation” (PID-D), “self-assessment about emotionality and rationality” (Emo. Self-Asses.) are relatively unimportant (values below 0.1), and the variable “overall prediction accuracy” (PAALL) is most important (highest value of the standardized discriminant), see the following table 44.

Table 44. *Standardized Canonical Discriminant Function Coefficients (3-Month Predictions)*

	Function
	1
COMSQ001	-.049
SQ001	-.027
SQ002	-.123
SQ003	**-.052**
SQ004	.081
SQ005	**.022**
SQ006	**-.027**
SQ008	.012
SQ007	**.025**
SQ009	-.044
Age Group	-.025
PID-D	**-.002**
PID-I	.043
Emo. Self-Asses.	**.036**
Skill Self-Asses.	**.073**
PAALL	**-1.500**
PA1W	.511
PA1M	**.713**
Survey Share	**.494**
Group No	**-.011**
Univ. Degree	-.024

The mean of the discriminant for both groups is shown by the following table (see Table 45); the average value for wrong predictions is 0.35 and for correct predictions -0.279, i.e. a moderate difference (possible values are from -1 to 1).

Table 45. *Functions at Group Centroids (3-Month Predictions)*

3-Month Predictions	Function 1
wrong	.350
correct	-.279

Unstandardised canonical discriminant functions evaluated at group means

The following table shows the results of the real and the predicted (using the discriminant function) assignment of groups (correct and wrong predictions). Only 63.6% of the original grouped cases were correctly classified. This is just little above the random value (50%). This shows that explanatory power of the existing variable assignment is rather limited (explains only 15% of the variability):

Table 46. *Classification Results (3-Month Predictions)*

		1-week predictions	Predicted Group Membership		Sum
			Wrong	correct	
Original	Frequency	wrong	1205	1067	2272
		correct	928	1344	2272
	%	wrong	53.0	47.0	100,0
		correct	40.8	59.2	100.0

a. 56.1% of original grouped cases correctly classified.

5.6.2.4 Discussion and interpretation of the discriminant analysis.

The discriminant analysis identified some variables with potential impact on the decision-making and predictive accuracy. The following table (see Table 47) provides an overview of the significant variables for the different prediction periods.

Table 47. *Determined Significant Variables*

	1-week predictions	**1-month predictions**	**3-month predictions**
Variables with significant differences between correct and wrong predictions	Commitment	Commitment	Group results
	Fundamental analysis	Financial ratios	Intuition
	Group results	Expert opinions	Market Sentiment
	Market Sentiment	Technical analysis	Expert opinions
	Age Group	PID-D	PID-D
	Survey Share	Skill Self-Asses.	Emo. Self-Asses.
	Group No	Univ. Degree	Skill Self-Asses.
	Univ. Degree		Survey Share
			Group No
Variables with significant influence on the distinctness (based on standardized discriminant analysis)	Commitment	Commitment	Group results
	Fundamental analysis	Financial ratios	Intuition
	Group results	Expert opinions	Market Sentiment
	Market Sentiment	Technical analysis	Expert opinions
	Age Group	PID-D	PID-D
		Skill Self-Asses.	Emo. Self-Asses.
		Group No.	Skill Self-Asses.
			Survey Share
			Group No.

There is no single variable that reaches significance level in all periods. However, there are a few variables that reach significance level in different periods e.g., “Commitment” (COMSQ001), “Group results” (SQ003), “Market Sentiment” (SQ006). However, not all variables turned out to be of significance. There are also a few variables that appear to be comparably less important in this context, since they did not reach significance level at any period, in particular “Use of company information” (SQ004), “News” (SQ008), and “Preference for intuition” (PID-I) (see Table 48).

Table 48. Overview *Variables Based on Discriminant Analysis Results*

	Variables with signif. differences between correct and wrong predictions			**Variables with significant influence on the distinctness**		
	1-week predictions	**1-month predictions**	**3-month predictions**	**1-week predictions**	**1-month predictions**	**3-month predictions**
Commitment (COMSQ001)	x	x		x	x	
Financial ratios (SQ001)		x			x	
Fundamental analysis (SQ002)	x			x		
Group results (SQ003)	x		x	x		x
Company (SQ004)						
Intuition (SQ005)			x			x
Market Sentiment (SQ006)	x		x	x		x
News (SQ008)						
Expert opinions (SQ007)		x	x		x	x
Technical analysis (SQ009)		x			x	
Age Group	x			x		
PID-D		x	x		x	x
PID-I						
Emo. Self-Asses.			x			x
Skill Self-Asses.		x	x		x	x
Survey Share	x		x			x
Group No	x				x	x
Univ. Degree	x	x				

5.6.3 Factor analysis.

In a factor analysis only numeric variables are allowed (Backhaus et al., 2016), therefore only a few variables of the data set could be introduced and some variables of the discriminant analysis could not be used (e.g., dichotomous variables such as “university degree” or ordinals in the factor analysis variables with three to five stages such as age group, group-membership).

Implementation of the factor analysis with the variables “overall predictive accuracy” (PAALL) and related sub-variables was not possible, presumably because of the high degree of correlation. Overall, this results in a limitation to six utilisable variables whose averages, standard deviation and sample size are shown in the following table:

Table 49. *Factor Analysis – Descriptive Statistics of Utilisable Variables*

	Mean	Standard-deviation	Analysis N
Commitment (COMSQ001)	2.66	.949	4518
Preference for Deliberation (PID-D)	3.712628498352256	.641551144315982	4518
Preference for Intuition (PID-I)	3.005634253110991	.466329548548778	4518
Emo. Self-Assessment	2.54	1.100	4518
Skill Self-Assessment	4.786	2.4117	4518
Overall Predictive Accuracy (PAALL)	.518155	.0715466	4518

The graph below (see Figure 19) shows the position of the variables spanned in two components: On component 1, the variables PID-I (negative), Emo. Self-Assessment and Skill Self-Assessment (both positive) show high values, while the variable “Overall Predictive Accuracy” (PAALL) shows high values on component 2. The variable “Commitment” (COMSQ001) is relatively insignificant (values close to 0 on both components).

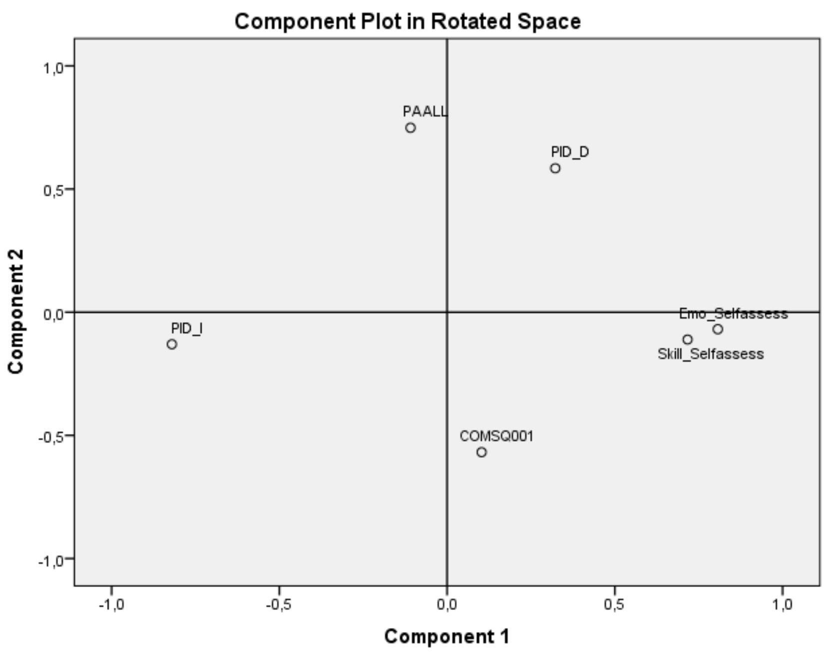


Figure 19: Component Plot in Rotated Space

The following coefficient matrix of component scores is based on the above graph (see Figure 19).

Table 50. *Factor Analysis – Component Score Coefficient Matrix*

	Component	
	1	2
Commitment (COMSQ001)	.036	-.049
Preference for Deliberation (PID-D)	.072	.150
Preference for Intuition (PID-I)	-.501	-.106
Emo. Self-Assessment	.372	-.089
Skill Self-Assessment	.159	.001
Overall Predictive Accuracy (PAALL)	.019	.684

Extraction Method: Principal Axis Factor Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

As a comparison of rotation methods a factor analysis was also performed using the Varimax rotation method (based on the correlation matrix). This is analogous to the Oblimin rotation method. The total variance explained using the Varimax rotation method was, for the first two components, approximately 53% (for details see the details in appendix “Quantitative Factor Analysis” on page 324). The rotation methods Varimax and Oblimin come to very similar values and component loadings.

The interpretation of the components is as follows: For component 1 the variables PID-I and Emotional Self-Assessment are the most relevant variables. However, this might not be very surprising, since both variables are aiming to measure the “intuitiveness” of the participants in a wider sense. Still, it is interesting that the PID-D variable, as a kind of “counterpart” of PID-I, does not follow this pattern.

5.7 Qualitative Interview Analysis

Overall, 59 people participated in the experiment. 25 participants were interviewed face to face, 29 participants by telephone, and three participants via email. Two of the participants dropped out during the experiment and refused to give an interview. Each interview typically took between 30 and 60 minutes. The interview data transcribed from all participants combined are in total a dataset of more than 80,000 words.

5.7.1 Coding conventions.

The following coding conventions were applied during the analysis of the interview data. The interview sessions were recorded and the audio files were completely transcribed. All interviews were imported to MAXQDA, where all coding of the interviews was conducted. The codes were clustered in several categories to allow a more systematic analysis. The categories used in the systematic qualitative interview analysis were inspired by Kuckartz' (2014) proposed spectrum of categories, but the categories have been selected and adjusted to fit the research question. Concretely the code system applied is a combination of deductive and inductive codes (Mayring, 2002, 2010), i.e., the code system was set up before the coding, but codes and subcodes were added during the coding process as well. For the coding and assignment of categories the following rules were applied: Initially all interviews were coded according to the respective questions from the structured interview questionnaire (see Appendix: Interview Questionnaire (Pilot Experiment) p. 295) with 23 codes, i.e., one code per question asked during the interviews (see also Figure 20). The interview documents were grouped according to the groups in the survey (AG, EDG, IG, NFG, and PG). In a second step thematic codes were applied and evaluative codes were added to the text. Most of the evaluative codes were created as sub-codes of other code categories. With this approach it is possible to code text snippets with more than one code. The codes are grouped in a hierarchical ranking according to the structure of the questions from the semi-structured interview outline. However, the coded segments might be statements from participants in the context of other segments as well. The classification followed the basic concept of interpretation and assessing meaning coherence in the context of the experiment conducted.

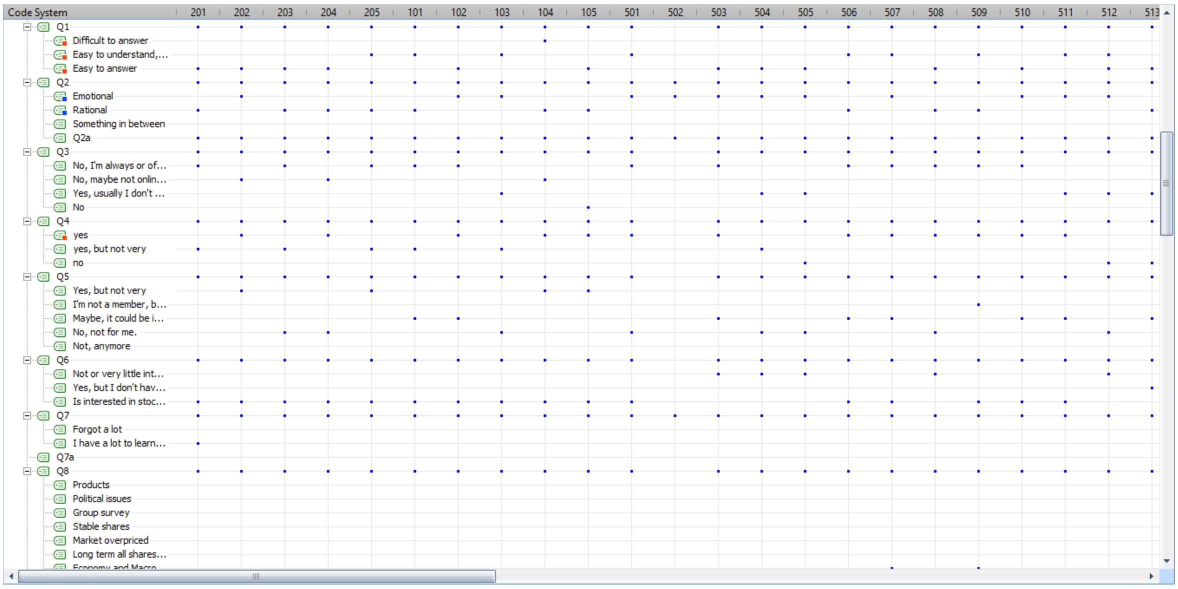


Figure 20: Screenshot of an Extract of the Project in MAXQDA Code Matrix Browser

5.7.2 Analysis of the interviews corresponding to the interview structure.

The following section contains the analysis of the responses from the participants with respect to the questions from the structured interview questionnaire. While the interview questionnaire was the orientation guideline for all interview situations, the answers were not necessarily always given in the same strict order. Accordingly, the responses from participants have been sorted to the corresponding question. The following section aims to provide an overview of typical responses in the interviews and to get a flavour of the world of thought, mind-set, and mentality of the participants. There is also some elaboration on the participants' opinion of the experiment design. The quotes are representative to provide a flavour of interviews.

Question 1:

The answers to the question: “Did you feel it was easy to answer the questions?” was typically that the questions were easy to understand in terms of clear formulation and not complicated, but were still not easy to answer for quite a few of the participants. Statements like “Yes, I understood the questions, but it was still not easy to answer, because prognosis is always difficult”. According to the answers given in the interviews it was easier for experts to answer the questions than for laypeople. Nevertheless, even some experts stated that it was not easy to answer. An expert stated “no, it was not easy and the reason for this is that I have only a commitment [an opinion] for a few instruments that are available at the market at a given point in time and it would be a coincidence, if that were one of the instruments that you picked for your survey and I have a dedicated opinion for that.”

Question 2:

The answers to the question, “Would you describe yourself rather as an emotional or a rational person?” provided some indication that rational behaviour is seen as a preferable character trait. Statements like “I would like to be rational and I always strive to be rational”, “emotions drive me more than I like. I always try to be more rational” or “over the years and as I've grown older I've learned to be more rational”, even though some people emphasize that it might depend on the situation and are not happy to describe themselves as simply emotional or rational. Still, many participants stated that in financial situations and with financial decisions they are more rational. “This is situational. There are situations where I'm completely emotional, but in general I was formed by political work and deliberative, at least used a deliberative approach to answering your questions”.
For the experts it is true that most, actually all except one of them, describe themselves generally as more rational people.

Remark: This finding appears to be particularly interesting, since most (almost all) participants answered during the survey that intuition and gut feeling was one of the bases for their decisions. It was by far the most frequent answer during the experiment. There is obviously a mismatch between the answers from the interviews, the answers from the online survey and to question 8.

<u>Question 3:</u>

The answers to the question: “Did you need to change your Internet usage in order to participate at the survey?” showed that they generally didn't have to change their Internet usage to participate in the survey, also the survey frequency – twice a week – was quite high. Most participants answered that they use the Internet and also check their emails on a regular basis anyway. Quite a few participants mentioned that they sometimes didn't use a PC, but filled in the survey on their smart phones while travelling or at home on the sofa. Some participants stated that they had to navigate to different Internet pages from those they would normally visit. Only very few participants stated that they don't use the Internet that often and that they had to go online on purpose in order to participate in the survey.

<u>Question 4:</u>

The answers to the question, “Are you member of a social community like Facebook or LinkedIn?” was dominated by the answer that people are members of at least one social community. Social communities seem to be a part of everyday life for most participants, but significance for day-to-day life is still limited “Yes, of course. I'm active passive user. I don't update my profile a lot, but I follow actively what others do.” or “Yes, I use it regularly, but that doesn't mean daily”. Even so, a number of participants said that they are members, but not very active or not very active any more. Statements like “[A member of] LinkedIn, but passive. If LinkedIn were a church, I would only go at Christmas” or “Yes, but not very active. I'm on XING and Facebook, but not very active there. I'm in and people write to me, but not more.”

Question 5:

The question, “Are you a member of a stock-market community like sharewise or marketocracy? If not, could you imagine becoming a member?” was answered by most people with “No.”. A number of participants from the lay groups were even unaware of the existence of such communities and how they work. They asked questions like: “What do you do there?” or “. . . if I knew what it was?”. A few people answered that they could imagine becoming members of such a special interest community, mainly to optimize their own investment decisions “I think it's interesting, definitely. [Could you imagine becoming a member?] Yes. I think exchange of information might take place easily, or if you have questions you turn to the online community and do not have to go to your bankers. Bankers also change very often nowadays and it's difficult build a bond of trust with them.” Another participant stated “Yes, I could imagine becoming a member? [What might be the argument for you to join such a community?] Simple interest in investments and to have a look at it. I'm more safety-oriented, because I once lost a lot, but realize now, that if you have a high level of orientation to safety you get no gains. Therefore, I try to compensate this by information. Therefore, in my own interest”. Still, most people don't think that this is a hot topic for them, and there are three main answers explaining why they don't want to join such a community. The first reason is “I don't have enough money” or “I haven't won the lottery”. The second reason is “I don't have enough time” or “Maybe, if I had enough time to concern myself with it”. A third, and most often mentioned, reason not to become a member is that people are simply not interested in the stock-market. “No, honestly I don't. That doesn't interest me so much. So I would say no.” or simply “No, that's not for me.”

For professionals / financial experts it appears to be slightly different. Financial experts are generally aware of the existence of such communities. Some of them have accounts to these communities and read there more or less regularly. “I'm a member of Wallstreet-Online and Ariva . . . I read there regularly” or “Not active. I have a few accounts in order to be able to read a few things, but I do not write there.” Some other professionals don't want to become members. “No, no I could not imagine. Because I think it is important always to keep one's distance from the herd. Even here in the office, I think it's important not to get carried away. Because it is not so easy when the market falls by 2%. Let's say I sometimes find it quite interesting to see what the mood is, but then also quickly close it down again.”. Some also doubt the quality of such communities: “No. Because it has no added value for me because the people who are members are not the people with whom I want to share my information and know-how.”.

Question 6:

The answers to the question: “Are you interested in the stock-market?” showed a somewhat mixed picture. There are quite a few lay people who are quite interested in the general economy and also the stock-market “Yes [How does this manifest itself?] Since I own a couple of investment fund shares and I also look at the indexes. I mean, the DAX and the other big ones. And sometimes there are also discussions with friends or something. This is of course an issue, even if you have no money to invest right now.” However, most lay participants mention that they are rather uninterested or not interested at all.

For professionals the answers were completely different. Almost all professionals mentioned that they are very interested in the markets and even see it as a hobby. Most of them trade shares themselves on a private account. “I trade stocks myself. That is probably the strongest argument. It's a hobby of mine. I read a lot that has to do with or could have to do with it.” or “It manifests itself first of all that I also invest my private money in stocks, that is also in single instruments and also that I inform myself just before I do that. This is on the one hand, of course, a purely financial investment, but also on the other hand interesting. In a sense it is also a hobby.”

Question 7:

The answers to the question: “How would you self-assess your knowledge about the stock market? (1, no knowledge; 10, expert)” indicated that participants are quite cautious about their skills. Most laypersons mentioned rather low values. Most like values 2 or 3.

Further analysis is needed and to check for correlations between a) Self-assessment knowledge and being deliberate or intuitive b) Self-assessment knowledge (see also discussion of Area of Discussion A11: People who think that they know more about the stock market are able to make better forecasts. on page 190).

Question 8:

As the answer to the question: “How did you make your decision?” most participants referred to their intuition and gut feeling. Typical statements are “Well, my gut feeling”, “Mostly intuition, if it were shares that somehow touched my area of interest then I either had some expert knowledge or more detailed knowledge from news in the media. And sometimes I've talked to experts who either have shares or are active in the area. In a round or two that was the case. They then also influenced my answers.” or “Mostly by intuition, intuition is of course influenced by what you have heard in the press and since there was also ThyssenKrupp in the survey, there was quite a lot in the press.” or following a further request to describe gut feeling many people referred to seasonal cycles e.g. “Partially how to assess seasons on the respective sectors or how I would assess the outlook for the sector. Yes, [I based my predictions] mainly on general bases... [like] production, purchasing behaviour.”, “I based my predictions on short-term considerations. I say it now with a very concrete example: when the survey began, the first survey and I went through all the shares there was HeidelbergCement among them, then it comes to mind that winter is coming and in winter the economy in the construction sector is weaker, so there is a tendency to falling prices.” or “Well, if I'm supposed to describe the share of a sports' equipment manufacturer, and it's characteristics; and it's winter or early winter, then I would rather say it is going to fall. Just like the construction industry.” or “Yes, my considerations. For example, you've asked about Heidelberg Cement, I can imagine that now in December, a cement company is simply selling less, so such considerations. And the news, like information on Siemens in the News which then of course influences opinions.”

Some people also referred to the media, the Internet and market indices as a basis for their decisions like “General news situation.” or “. . . first the Internet to view the stock prices . . . and then the results of the weeks before.”

The professionals still referred to intuition and gut feeling. They gave answers like “Frequently market climate and intuition and gut feeling, probably even more than valuation, although I know the valuation and ratings of companies that I do not cover as well.” But it seems that gut feeling is not the same for professionals and some lay people, together with a similar initial classification, described a quite different decision approach. When asked to describe intuition and gut feeling they answered “Has a lot to do with the development of the stock in the last few days and how I generally assess the market. So for example, I guess the market is not so great, and the stock previously went very well, then I guess it's not so good, it's probably going down. For example.” or “Yes, more like the general market sentiment, the news flow, macro but also micro, so to speak, and how I perceive it, so that's not carefully analysed but rather the current mood.”

Question 9:

As answer to the question: “Did you prepare for the survey rounds? If yes, how?” most participants stated that they didn’t prepare for the survey. Some stated that they did some research in the beginning. “No, actually not. When it came to the first time, I started to research, but I did not make any preparation for it, no.” or “I did so when you sent it to me, but I didn’t really prepare actually, so maybe I had a quick 5 minutes”. Some participants decided not to prepare by purpose: “No, intentionally not . . . I deliberately didn’t bother about it any more, because then it would not really reflect how I really am.”

Most of the professional experts also stated that they didn’t prepare beforehand and just did some research when they started to fill in the questionnaire. Still, their way of dealing with the survey appears to be more sophisticated: “Yes, passive. Yes and no. Well, as to the companies that I do not actively cover, like HeidelCement and ThyssenKrupp, I nevertheless know the general news flow because of my work. The indirect the preconditions were there, but I didn’t sit down specially and google something, no nothing like that.” or “Well, I also checked at the beginning whether there were any outliers. I do not know maybe I also looked again during the survey to see whether something serious had changed, but then during the survey rounds it's more like the general market sentiment is relevant, and what can be deduced from it. I don’t think that I'm the expert on individual companies and what is special about them and whether the products are good and if that then also affects performance in the short term.”

Question 10:

The answers to the question: “Did you use external sources for the experiment? If yes, which ones?” provided a varied picture. Some participants stated that they didn’t use any external sources e.g. “No, I didn’t” or “No external sources”. Some stated that they hardly used any external sources “No, not really. Maybe only once or twice, there was external media for me, radio and television, which presented a report about the ones that have problems and are probably going down. I believe that was ThyssenKrupp and I thought to myself, they can’t do much. Otherwise, I mostly did what I thought [best].” If participants stated that they used external sources, the sources were mainly newspapers or Internet portals “Well I googled. It was mostly Finanzen.net. I just googled share prices and saw how the share price was in the past, and clearly at the end you also have to estimate. . . . I mean on the Internet you will find that in some cases, and even more, the company's key figures.” or “Partially Finanzen.net, otherwise information I came across by chance”. The mentioned online sources were Finanzen.net, Onvista, Börse.de, Boerse-Online, and Finanznachrichten, but also the websites of Online-Brokers (CortalConsors and Comdirekt). The answers from financial professionals showed that they generally have slightly different access to financial information “Yes, Onvista and Bloomberg. In each case the indexes – technical analysis would probably be saying too much. I know the analysts' opinions.” “Yes, Index Information Services and Markets Pages [Which ones?] Google, Yahoo and Onvista mainly.” An interesting fact might be that they used mainly similar information sources to lay people and if they used professional information services like Bloomberg they mainly checked the information, like price indexes, that are also available free on the Internet. A major difference to lay people might be the frequency of usage and the interaction with other financial professionals in the decision-making process. “Yes, you spoke, in the case of one company or another which was part of the survey, with a colleague that has greater know-how for that case.”

Question 11:

The answers to the question: “Did you ever buy shares? Have you bought some of those used in the experiment?” again showed a quite different picture for professionals and lay people. All professional participants in the survey stated that they actively trade stocks. Typical statements from the professionals are “Yes, and even some of those in the experiment” or “Yes, all from the experiment. Many of them long and short.” The question, if they had a somewhat closer relation with the stock and if it was more easy to make an estimate was answered like this "Yes, you already dealt with it, if it was 3-4 years ago or just a few weeks before, ultimately you have a stronger feeling for it." or "Yes, if you have some trade experience you have usually acquired more knowledge about it."

Many lay people answered that they don't buy or trade shares or just as an investment fund "Not directly. I had shares in an equity fund" or "Only funds. So not really shares, but only funds". The few lay participants that had done so usually also stated that they rarely follow-up the issue: "Yes, I did play the stock market and I once inherited shares, but I didn't really follow them and then eventually sold them." or "Yes, I once had shares, but they are all gone." Just a few individual lay people seem to be really interested in single stock investments "Yes. [Also some which were mentioned in the experiment?] Yes, Siemens and ThyssenKrupp. [Then you had already some relationship with them?] Yes, although I did not own them at the time of the experiment".

Question 12:

The reactions to the question: "Are you (EDG/IG) / Would you have been (NFG) interested in the Group results?" were quite different between the groups. The groups with feedback (EDG/IG) stated that they had a look at it, like "This was interesting for me because then I saw how I assessed it and how the others had estimated and that was really interesting, yes." or "Yes, from time to time, yes", but often just with limited interest in the group results e.g. "I looked at the first three times, and then it was somehow too stupid. Was always the same", "Rarely actually, maybe I looked at it two or three times, but not more often actually." or "I skimmed through it. That means I checked what the others say, but I didn't align my judgement or estimates for the next survey according to it".

Some people were somewhat more interested in the results. "It was very interesting for me, I noticed that I was in-line with the trend with some shares and clearly against the trend with others, but I did not let it affect me. I found it interesting to see what the swarm is thinking.” Some even tried to get some information about themselves out of it “Well, that did not affect my opinion for the next time, but I was in so far interested as I was interested to see whether I'm a mediocrity or whether I'm doing what the majority does, or whether I do it intuitively. Whether I'm such a majority person. Whether I'm such a herd animal or not, but it was more psychologically interesting for me, but not that I think that my share would rise if I sat on the horse the majority bets on.” Generally the tenor was that the group results were of limited interest and the participants felt that the group didn’t impact them a lot or if so just to confirm their own beliefs: only a few individuals stated that the group had an impact to their own decisions “Yes. I used it for orientation”. It can be noted that only participants who rated their own knowledge about the stock market as quite low (e.g. 2-3) used the group results for orientation.

The participants from the group without group feedback generally mentioned more interest in the group results, and many participants assumed that the group results might have impacted their own decision-making.

Question 13:

The answers to the question: “Do you think the e-Delphi-experiment / the group results influenced your decisions?” confirmed the answers from question 12 in the sense that most participants stated that they were not influenced in the decision-making process by the group results. “Nope. I didn’t rely on them [the group results]. I have my own opinion” or “No, if I had looked at it more often, then maybe yes, but since I looked at it only 2 or 3 times and skimmed over it briefly and had a look out of interest, I would say that it has not influenced me.” Some participants assume that it might have unconsciously influenced them. “Not consciously, subconsciously certainly.” “No, I do not think so. If then only very subconsciously and I can’t tell you.”

For some participants the group results were also some reassurance for their own beliefs “I do not think so. Perhaps in the sense that I was a bit confident with my own assessment." or "No, it did not influence my decisions. I just looked at it and saw that I'm not quite wrong, so I can go on in that direction."

Even so, a few participants stated that the group results had an impact to their own decision-making. "Yes, yes, because I still needed somehow something because I looked on the Internet and I was able to find some information, but I was still relatively uncertain. And at that time the group was a help and a guideline for me." or "Yes. I think so. [How so?] Because in some cases I was not quite sure and then I orientated myself by the results."

The participants from the professional group stated that they didn't follow the group results at all. "I hope not." or "No. I'm autonomous with my opinion."

Question 14:

The answers to the question: "You changed your decision [X times from Y to Z] in round 2; why?" showed that a lot of participants were quite unaware of their changes and their frequency. While lay people in particular assumed and stated that their decision-making rationale was quite stable during the three months of the survey (e.g. season or economic cycle). "Really? I didn't even notice it. That's quite a lot.", "Maybe I had more time to think about it. On Sunday, or at the weekend I have more time to reflect about what happened during the week. That would be an explanation for me." or "No, well honestly not. Honestly I'd have thought I was pretty much the same all the time. Because the survey was, yes, in the winter, in the autumn-winter and so I actually thought that I had always made quite the same decision. So I cannot explain why I changed my opinion". Some participants who changed quite frequently contrary to the group decisions were also asked if they used the group as a contra indicator, but all denied that and insisted that their own opinion was more important to them like "I wouldn't have been aware of that. Of course I know the option that you don't go with the crowd, but that was not my conscious decision. It was more that I had the feeling that it was going in that direction. For the evaluation, the group was less decisive for me, but rather the certainty with which I rate it."

Question 15:

The answers to the question: “You mentioned that your decision-making is based on [XY], and you did/did not change [changes in direction] during the experiment; why?”. A few topics were mentioned across all participants, e.g. gut feeling and news flow. Typical statements are like “In the beginning my decisions were based on gut feeling, market sentiment and news flow, but later on in the survey just gut feeling and news flow”. However, a closer look at the differences between the top and worst predictors reveals some differences. None of the worst 6 predictors relies on expert opinions, but 3 of the top 6 predictors mentioned expert opinions as an influence on or basis for their decisions. The use of technical analysis seems similar: it was not mentioned by the worst 6 predictors, but at least 2 of the top 6 predictors mentioned that their decisions were at least partially based on technical analysis. Generally it appeared that the top predictors tend to question the reliability of their sources more than the weak predictors and they are more willing to adapt their decision-making approach or at least the source of information. As one of the top predictors puts it “I have to know the expert to get a picture and to trust his opinion”. While the poor predictors rarely mentioned any substantive discussion of the quality of their sources and basic principles.

Question 16:

The answers to the question: “Were you influenced by the group?” were very homogeneous and everyone stated that the group had no influence on their approach to decision-making. “Not cautious. I know you can't be neutral, but I've tried to look at it neutrally. I didn't say, “I go with the group then I'm right” or “I looked at the group results and then I agree with them, then I'm right or I did not say I am against it because it's cool. I looked at the results, but then I also tried to put it aside. So I didn't view the group evaluation immediately before filling out the surveys, nothing like that.”

Question 17:

The answers to the question: “Do you think you gained new expertise or knowledge during this experiment?” showed that the participants were quite cautious about that. Some people stated that they didn’t gain any knowledge. “No, not at all.” Some of them think that they gained a little knowledge “Um, well, yes. Not a lot, but still a little bit.” For a few participants the three-month period was more of a learning experience. “Yes, of course, because everything you deal with just deepens your knowledge. It's like in journalism in the end, there you also know just a little about the matter at the beginning and if you then investigate the matter you know more. Expert is perhaps a bit exaggerated, but you will soon know more. And so I found that in that case too, so [it was] really exciting.” or “Yes, well at least my interest has increased enough to have a closer look at it.”

One of the participants stated that he didn’t learn anything because he was annoyed by the high frequency and the fact that there were always the same questions. “No, because it was always so stressful that every week there were the same shares.” So generally the lay people stated that they had gained no or only limited expertise because of or during the experiment, referring mainly to a lack of time and interest in the subject matter. “No, but I should have prepared myself better for it and maybe talked to an expert.”

For professionals it was much the same. “Nothing ground-breaking, but you always learn a little bit” or “If so, then only nuances. I always follow ThyssenKrupp and Siemens very closely. To a certain extent yes, but not very strongly.”

Question 18:

The reactions to the question: “Do you care more about news now, in particular news about the companies in the survey?”. Many participants stated that they didn’t care more about news. “No, I don’t think so”, but most participants confirmed that they cared more about the news flow related to the companies. “Yes, yes, because when I heard something during that time, I listened more carefully, because I know the next question about it will probably follow. So you are more likely to listen more closely.” or “Yes, probably.. . . No, I mean the ThyssenKrupp example is of course a great one. I don't know if you knew when you chose it that it was going that way. I was not aware of the situation until the middle of the study. This is of course very exciting in terms of who's buying whom and so on. An exciting example. So let's say this, if I had some dosh I would buy Thyssen now.”

Question 19:

The reactions to the question: “What do you think about the usability of the web survey tool?” generally confirmed, that the tool was appropriate for such a purpose from the point of view of the participants. “I thought it was cool. What I've always been waiting for. I was fun for me and it also worked very well.” or “I thought that was not bad. It was well designed and clear. I liked that.”

Quite a few highlighted the fact that usability was a crucial point in their participation: “Well, that was uncomplicated to use and the reminder by email and the link that took you there directly were good, because otherwise I probably would not have done it.”

Some participants emphasized that optimization for mobile access would be a nice feature. “I think that's actually quite good. Maybe you could highlight where you are currently a little more, in terms of “it is the umpteenth poll now”. It always looks a bit similar and it could work even better with mobile devices. I often filled it in on the phone in the evening, because I was not in front of the computer, so that I could still can make it in time to enter the estimates. Some optimization is a possibility. But basically I think it is good in terms of user interface.”

Question 20:

The answers to the question: “What would you like be changed about the survey?” showed that the participants didn't like the rhythm and monotony of the questions. “More varied questions. I think I had already said that before. It was kind of one-sided and, as I said, my opinion of it has not really changed. It was, I think, twice every week and since there were always the same questions it was boring. I would say.” or “Of course there is something. So, what would I do differently? But then you probably wouldn't achieve your objectives. You do that every week, which I also found a bit of a pain or you could take other shares . . .” A few of the participants used this question to highlight how they appreciated that the questionnaire was short. “I think it was nice, inasmuch as it was short and precise, so that you didn't have to fill in a 4 page long questionnaire. But, that I could take in a whole page at a glance and then I could click on it. That's always an aspect of usability. In this respect, I found it good. I also found the interface was presented in a modern way. So I cannot say anything negative.”

Quite frequent was also the suggestion that it would be nice to see the current share price, maybe even with a small chart. Another quite frequent recommendation was to show or send a summary of the own results. “Add the possibility to review in the end everything what you have entered, on scrollable page.”

One participant felt that just interacting with a computer is not sufficient: “Maybe to interrupt this automatic process from time to time and to provide personal feedback or facilitate a short exchange within the group in some way half way through or after the first third, and after the second third. . . . Maybe a forum or something similar. That you are not alone in front of the automat.” Another participant wished there had been some more introduction to the topic and to brush up their knowledge about financial markets. Yet another suggested improving readability. “I would make the overview of the group results bigger. The one you sent after each round. They were in a quite small font. Apart from that I thought it was good. It was clear and I never had any service problems."

One of the professional/expert participants made the following suggestion: "I would have liked 4 values and one from DAX and MDAX so that I could pick and choose myself, so that I could really reach a high level of commitment."

Question 21:

The answers to the question: "You didn't enter a price target [in X of the cases], why?" can be divided into three groups, or maybe even just two groups, because one of the groups always entered the price targets and therefore didn't have to answer this question. The second group is basically the one that sometimes forgot to enter a price target. "That was probably an oversight. There was no bad intention in any case." or "Maybe because it was on my mobile and an input error." The third group didn't want to enter a price target and figured out that price target is not a mandatory field. This third group was actually the smallest group (only 3 participants).

Question 22:

The reactions to the question: "Do you think it's easier to enter a concrete price (in Euro) as target price or a percentage change?" were quite different for professionals and for lay people. While the professional stated that it didn't matter to them or that they prefer percentage. "No, quite the opposite. People are indeed different. I always watch using %, otherwise I would have had to convert the value." Only one analyst emphasized that concrete prices would be better for him. "I think yes, so on the basis that I am quite familiar with the prices. But for people who do not know the prices? But if you have to predict the prices it somehow requires you to have a rough idea of the price. So I think I would have made somewhat more logical predictions with the Euro amount. Because you sometimes entered round values like 10 or 15 %, somehow you do not give 7.5 % or values like that.

But if you try to think like a trader, you are talking about absolute price amounts. The same with ratios and figures. . . . If you expected a PE you would not be likely to consider a 10% upside, but rather the stock could be at 60 or 70. I probably would have found it a bit easier."

For lay people it was quite different story. Some participants highlighted the fact that they didn't know the concrete share price and in this case the percentage figure is much better. "No, I think the percentage is quite good, because I don't necessarily know the absolute amounts, because I do not follow the prices. The percentage is good, because it is not as important in that case as the numbers are high now." On the other hand percentage is not the preferred version for some lay people. "Yes, percentage is always a little bit abstract for me." or "Yes, it would have been easier for me. I think so. I would think Euro amounts are better for me than percentages. [Why?] Because I can think about it a more concrete manner."

Nevertheless, for many lay people as well it appears to be no problem to use percent or concrete amounts. "No, I do not think so. It was okay with the percentages. It would have made no difference to me whether you specify in Euros or in percent"

Question 23:

The answers to the question: "Any further comments or suggestions?" were also quite different. While a lot of the participants said that they didn't have any further comments, a lot of participants provided a wide range of feedback and were interested in getting the results of the survey. "I would just like to get a final result of the study." One participant suggested supplementing the survey with a forum.

5.7.3 Comparative analysis of best and worst predictors.

The following table (see Table 51) shows an outline of noteworthy parts of a comparative analysis (also including frequencies) of the interviews of the best six and worst six predictors—based on their overall predictive accuracy—of all the participants. Please refer to table 130 in the appendix on page 305 for a more comprehensive overview of the comparative analysis. The comparison included all codes which appear to have some potential to add meaning and might allow differentiation between the participants.

Table 51. *Comparative Analysis of Best and Worst Predictors (Code Matrix)*

	Participant Interview Analysis (Frequency of Answers)											
Participant ID	204	503	511	516	101	604	617	507	520	2	515	521
Rank of Participant (Based on Overall Prediction Accuracy)	1	2	3	4	5	6	54	55	56	57	58	59
Interview Questions												
Q1												
Easy to understand, but difficult to answer			1	1	1			1		1	1	
Easy to answer	1	1				1	1		1			1
Q2												
Emotional		1	1					1	1		1	
Rational	1			1	1	1	1			1		1
I try to be rational/Emotional is not good						1						
Q3												
No, I'm always or often online		1		1	1	1	1	1	1	1	1	
No, maybe not online but mobile.	1											
Yes, usually I don't check daily			1									1
Q6												
Not or very little interested in the stock market		1									1	2
Is interested in stock market	1		1	1	1	1	1	1	1	1		
Q8												
Economy, Politics and Macro						1		1				
Business Cycles/Seasons						1					1	1
News and Media		1					1	1	1	1	1	2
Technical Analysis		1	1	1								
Company Information or Analysis	1		1		1							
Market sentiment	1			1	2					2		
Opinions from others (incl. analysts)	1			1								
Peers, Friends, Experts					1							

Intuition/Gut feeling/Smart Guess			1		1		1		1	1		
Q10												
No							1	1				
no, but...												1
Yes, Newspapers, Magazines						1				1		
Maybe speaking with experts	1											
Yes, Internet Portals		1	1	1		1			1	1	1	
I checked some facts and figures					1							
Q14/15												
Different Mood/Weekend more time						1	1					
Company specific issues						1	1					
Group communication/group influence										1		
Market sentiment / Political Issues			3									
Experts	1		1	1								
Intuition/Gut Feeling		2		1				1		1	1	
No cautious decisions/by chance				1						1	1	
Technical Analysis	1		1									
News-flow		1	1				1	1				2
Q18												
No, I don't care or don't pay attention									1			
No, I already had a strong background before.					1							
Yes, I think so.	1	1	2	1		1	1	1			1	
Q22												
Wouldn't ask any of these questions	1											
Euro more easy									1	1	1	
% is better		1	1	1		1	1	1				1
Doesn't matter to me.					1							

In summary, the interview analysis of the six best and worst predictors reveals that all participants understood the questions and that it was equally difficult for good and poor predictors to answer (three of each mentioned that it was difficult). There was also no perceivable difference in the use of the Internet. All participants use the Internet at least on a daily basis.

There are differences between top and worst predictors in other domains. It seems that for good predictors the ideal is to make “rational” decisions. One of the top predictors even mentioned explicitly that he tries to be rational because “emotions are not good”. All but one top predictor mentioned that they are interested in the stock market, while at least two of the poor predictors mentioned that they are not or only very little interested in this topic. A particularly striking difference is in the bases on which they make decisions: while poor predictors mentioned that they rely on news (and gut feeling) to a large extent, there was a much more differentiated picture for the top predictors (this is reflected in particular in the answers to question 8 and questions 14/15). Top predictors also emphasized slightly more that they were more sensitive to news related to companies in the experiment (Q18).

Another difference between top and poor predictors is that none of the top predictors thought that it was more easy to provide concrete predictions in Euro rather than a change in percent, while three of the poor predictors mentioned, that they would have preferred to enter a Euro amount rather than a percentage value.

5.8 Triangulation and Areas of Discussion

The following section provides a comprehensive overview, examination and discussion of various aspects using the data from the main experiment. Whenever it appeared to be appropriate, a triangulation of quantitative and qualitative data was conducted to reappraise and refine the findings.

5.8.1 Discussion and analysis of the experiment data.

The literature review, pilot run and main experiment revealed several areas of potential interest for further analysis. In total, 21 areas were discussed in the following analysis. Whenever in the following text a Chi-Square test was conducted, it means that it was performed with the Chi-Square tool created by Preacher[1] (2001) with one degree of freedom and a level of significance (p-value) of 0.05 (5%) or with the functions implemented in SPSS, if not explicitly stated differently. The discussed areas are as follows:

- A1 A lay person may be better at predicting short term (1 week) than a professional financial analyst, but over a longer period the analysis models of an analyst will lead to better results.
- A2 There is an improvement predictive accuracy resulting from feedback from an e-Delphi group.
- A3 Predictive quality improves over time as people learn about the shares.
- A4 Rational people are better at financial decision-making compared with intuitive people.
- A5 Educational level has a huge impact on the ability to predict stock prices
- A6 Female are better at determining market sentiment and perform better with short term predictions.

1 Preacher, K. J. (2001, April). Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Available from http://quantpsy.org.

- A7 Financial analysts are consistently over optimistic in their forecasts of the shares covered.
- A8 Life experience and age have an influence on stock price predictions. Older people are more risk averse.
- A9 Analysts are better than lay people in bull markets, but lose that advantage in bear markets.
- A10 People who are interested in the stock market are able to provide better predictions.
- A11 People who think that they know more about the stock market are able to make better forecasts.
- A12 Predictions with a higher level of confidence are generally better than predictions with low confidence.
- A13 When people express a higher upside or downside (in terms of price to target price difference) the predictions are better.
- A14 Predictions based on fundamental (or technical?) analysis are superior to intuitive predictions.
- A15 People who base their predictions on several decision-making approaches and/or information sources are better than those who decide based on fewer approaches/sources.
- A16 People are better at predicting steady upward or downward trends than changes of direction.
- A17 Certain individuals are especially good at predicting, as compared to the average of other members of a given group, of which they are members.
- A18 Predictions of well-known shares, such as Adidas, by lay people are better than their predictions of lesser known (to the general public) shares.

- A19 Some lay people regularly outperform others in a given lay group of which they are members.
- A20 The same lay people who are good at short term predictions are also good at longer term predictions, as compared to the group average.
- A21 If a good predictor is defined by having a higher value of number of correct predictions divided by number of incorrect predictions, then what are the characteristics of these good predictors, as found from the questionnaire results?

Area of Discussion A1: A lay person may better at predicting short term (1 week) than a professional financial analyst, but over a longer period the analysis models of an analyst will lead to better results.

A rationale for the assumption that lay people may better at predicting short term (1 week) than a professional financial analyst, but over a longer period the analysis models from an analyst (see also 2.2 Equity Research) will lead to better results might be that lay people may use common sense to assess market sentiment and subsequently market sentiment influences share price development in the short term. The analysis of data from the main run indicated no significant difference in predictive accuracy for the 1-week (Chi-square 0.19, *p-value*=0.66) and the 3-month (Chi-square 0.009, *p-value*=0.92) share price predictions (see Table 52 and Table 53). Furthermore, the accuracy of the 1-month predictions of the expert groups was considerably higher (Chi-square 7.2, *p-value*=0.007). Another interesting indication can be derived from the discriminant analysis (see Table 47), where the variable "use of fundamental analysis" is significant for the one-week predictions, but not for the longer periods (one- and three-month predictions).

Table 52. *Comparison of Correct Answers in the Main Experiment*

	1-Week (correct answers)	1-Week (wrong answers)	1-Month (correct answers)	1-Month (wrong answers)	3-Month (correct answers)	3-Month (wrong answers)
Lay Groups	50%	50%	46%	54%	60%	40%
Expert Groups	48%	52%	59%	41%	60%	40%

Table 53. *Absolute Number of Correct and Wrong Answers in the Main Experiment*

		Lay Groups	Expert Groups
1-Week	Correct	140	78
	Wrong	140	85
	Excluded	20	37
1-Month	Correct	126	101
	Wrong	148	70
	Excluded	26	29
3-Month	Correct	171	103
	Wrong	115	68
	Excluded	14	29

Overall, it can be concluded that assumption A1 is not supported by the data generated in the experiment. There was no significant difference in their predictive accuracy for two of the forecast periods in the main run. Additionally, assumption A1 does not provide a rationale for the good expert predictions for the 1-month period.

Area of Discussion A2: An improvement in predictive accuracy results from feedback from an e-Delphi group.

Group learning is a basic feature often attributed to Delphi and e-Delphi processes (e.g., Dalkey, 1969; Rowe & Wright, 1999). Consequently, the result expected from the experiment could be that the predictions in a 2nd e-Delphi round, after the feedback from the group, would tend to result in more accurate predictions. The data gathered in the experiment provides two benchmarks to test A2. The first benchmark is a comparison of the accuracy of the first (see also Table 28) and second (see Table 29) e-Delphi-Rounds within the groups with feedback loop (EDG, IG, and PG). Another benchmark is the comparison between the groups with (EDG, IG, and PG) and without feedback loop (AG, NFG) (see Table 54).

Table 54. *Influence of the Feedback Loop (e-Delphi Round 1 to Round 2)*

		1st e-Delphi Rounds			2nd e-Delphi Rounds		
		All Groups	Groups with Feedback	Groups without Feedback	All Groups	Groups with Feedback	Groups without Feedback
1-Week	Correct	135	56	50	137	57	55
	Wrong	125	74	36	136	80	35
	Excluded	40	20	14	27	13	10
1-Month	Correct	144	67	50	140	66	44
	Wrong	115	57	41	136	73	47
	Excluded	41	26	9	24	11	9
3-Month	Correct	158	79	50	176	95	50
	Wrong	105	50	40	108	50	43
	Excluded	37	21	10	16	5	7
1-Week	Correct	52%	43%	58%	50%	42%	61%
	Wrong	48%	57%	42%	50%	58%	39%
1-Month	Correct	56%	54%	55%	51%	47%	48%
	Wrong	44%	46%	45%	49%	53%	52%
3-Month	Correct	60%	61%	56%	62%	66%	54%
	Wrong	40%	39%	44%	38%	34%	46%
Overall	Correct	56%	53%	56%	54%	52%	54%
	Wrong	44%	47%	44%	46%	48%	46%

The overall accuracy of the groups in the 1st e-Delphi rounds was about 56% correct answers and the accuracy of the groups with feedback loop in the 1st e-Delphi rounds about 53% correct answers. The overall accuracy of the groups in the 2nd e-Delphi rounds was about 54% correct answers and the accuracy of the groups with feedback loop about 52% correct predictions. Thus A2 is not supported by the data since the overall accuracy even decreased slightly from the 1st e-Delphi round to the 2nd e-Delphi round.

Table 55. *Changes from e-Delphi Round 1 to Round 2*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	6	-2	1	-1	2	-4
	Wrong	-3	0	2	2	4	6
	Excluded	-3	2	-3	-1	-6	-2
1-Month	Correct	-3	0	-7	-3	6	3
	Wrong	5	-1	10	1	7	-1
	Excluded	-2	1	-3	2	-13	-2
3-Month	Correct	1	2	4	-1	10	2
	Wrong	0	-1	1	3	0	0
	Excluded	-1	-1	-5	-2	-10	-2
1-Week	Accuracy Change	10.44%	-2.34%	-0.86%	-3.37%	-0.48%	-11.56%
1-Month	Accuracy Change	-9.15%	0.93%	-19.40%	-4.17%	-6.85%	3.85%
3-Month	Accuracy Change	1.22%	2.64%	3.76%	-4.65%	6.33%	1.48%

The comparison between the groups with feedback and groups without feedback loop also indicates no advantage from the e-Delphi survey for the predictive accuracy of the groups with e-Delphi feedback loop. The accuracy of non-feedback groups is slightly higher in the 1st e-Delphi round (56% correct answers) and the 2nd e-Delphi round (56% correct answers) compared with the groups with feedback loop between both rounds. Groups with feedback loop had only about 53% correct answers in the 1st e-Delphi rounds and 52% correct answers in the 2nd e-Delphi round. Again this data does not support A2. There was no measurable improvement; on the contrary there was overall a slight (non-significant) underperformance by groups with feedback loops.

Area of Discussion A3: Predictive quality improves over time as people learn about the shares.

When people engage in something it is usually a side effect that they get more and more familiar with the topic. A3 basically assumes that a learning effect occurs during these periods and predictive quality improves (see 2.3.4 Structured group decision making. and 2.4.2 E-Delphi.). This assumption can be tested with the data and a comparison of the predictive accuracy of the predictions provided by participants during the first half (weeks 1 – 5) and the second half (weeks 6 – 10) of the experiment. If there is a learning effect it might be even stronger, when people only have a little initial knowledge about the topic. Accordingly, the lay groups should reveal the highest improvement in predictive quality.

Table 56. *Results 1st Half (Week 1 to 5)*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	19	14	20	30	11	23
	Wrong	18	33	30	19	24	20
	Excluded	13	3	0	1	15	7
1-Month	Correct	17	7	21	16	31	20
	Wrong	24	40	29	31	7	23
	Excluded	9	3	0	3	12	7
3-Month	Correct	11	33	22	20	32	20
	Wrong	28	15	28	27	4	23
	Excluded	11	2	0	3	14	7
1-Week	Correct	51%	30%	40%	61%	31%	53%
	Wrong	49%	70%	60%	39%	69%	47%
1-Month	Correct	41%	15%	42%	34%	82%	47%
	Wrong	59%	85%	58%	66%	18%	53%
3-Month	Correct	28%	69%	44%	43%	89%	47%
	Wrong	72%	31%	56%	57%	11%	53%

The data from the main run indicated that there was an improvement in predictive quality from the first half compared with the second half of the experiment. While the predictions of weeks 1- 5 turned out to be correct in 46% of the cases (see Table 56) it increased to a predictive accuracy of 64% of the predictions provided during the weeks 6 – 10 (see Table 57). Overall 367 predictions in weeks 1 - 5 turned out to be correct and 423 wrong and in contrast 523 predictions in weeks 6 -10 turned out to be correct and 302 wrong. This is a highly significant result (Chi-square: 46.802, *p*-value<*.001*).

Table 57. *Results 2nd Half (Week 6 to 10)*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Correct	29	28	21	27	19	31
	Wrong	15	21	18	19	28	16
	Excluded	6	1	11	4	3	3
1-Month	Correct	28	35	14	33	25	37
	Wrong	19	13	21	14	20	10
	Excluded	3	2	15	3	5	3
3-Month	Correct	30	37	20	39	30	40
	Wrong	18	12	23	10	18	7
	Excluded	2	1	7	1	2	3
1-Week	Correct	66%	57%	54%	59%	40%	66%
	Wrong	34%	43%	46%	41%	60%	34%
1-Month	Correct	60%	73%	40%	70%	56%	79%
	Wrong	40%	27%	60%	30%	44%	21%
3-Month	Correct	63%	76%	47%	80%	63%	85%
	Wrong	38%	24%	53%	20%	38%	15%

Furthermore, a comparison of the improvement of the lay groups and the expert groups does indeed show a considerably higher improvement in predictive quality. There was no significant difference in predictive quality during the first half of the experiment. The lay people with 42% Correct during the first half of the experiment were slightly less accurate than the experts with 43% Correct. In the second half of the experiment the predictive quality of the lay people was 63% accurate predictions and only 57% of the predictions by the experts proved to be correct.

In summary, the data from the main experiment supports A3 and it appears that there is a measurable learning effect in the participants during the period of the experiment.

Area of Discussion A4: Rational people are better at financial decision-making compared with intuitive people.

When it comes to financial decision-making, such as predictions of stock price movements, it might be conceivable that rational people have an advantage (see also 2.3.1 Financial decision-making). There are two variables in the data set which might reflect the rationality of the participants. First there is the Self-Assessment by the participants as taken from the interview data, coded with the categories “emotional”, “rather emotional”, “rather rational”, and “rational”. The second variable is the so-called PID scale value.

A box plot (Figure 21) of the interview data and overall predictive accuracy (PAALL) indicates that all four groups are at about the same level in terms of decision-making and accuracy. This indicates that A4 is not supported by the data gathered in the experiments and self-assessments from the interviews.

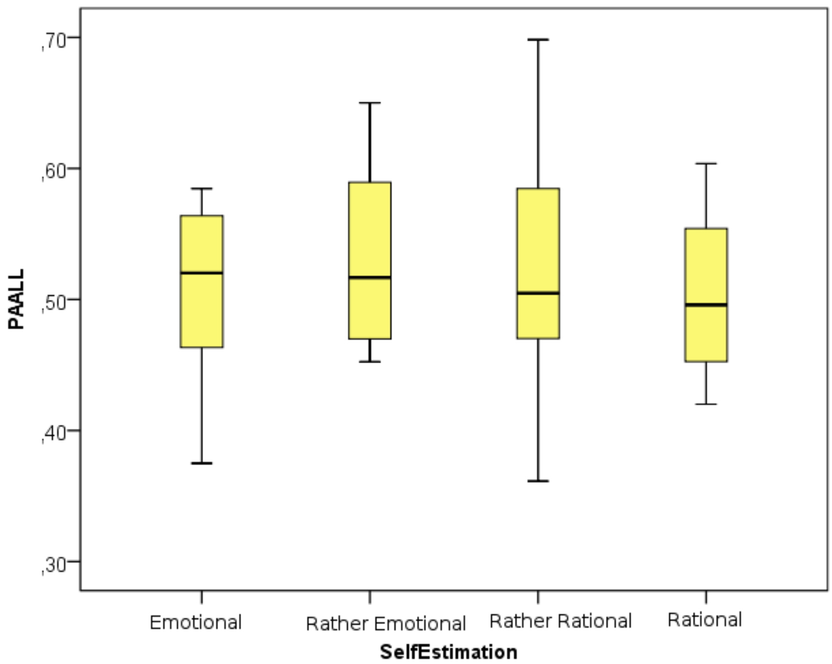


Figure 21. Self-Assessment Rationality/Emotionality and PAALL (Box Plot)

The analysis of the PID scale data might be even more informative (Endress & Gear, 2015). In order to measure this value all participants in the main experiment (*N*=59) were asked to complete a questionnaire to evaluate their preference for deliberation or intuitive decision-making (PID scale). The questionnaire was developed by Cornelia Betsch (Betsch, 2004; Schunk & Betsch, 2006). She presented in her thesis an assessment tool, the PID inventory, to determine the preference of people for deliberative or intuitive decision-making. The PID scale provides four categories to group people according to their preferences. There are people with a preference for deliberative decisions (PID-D), people with a preference for intuitive decisions (PID-I), and people with situationally varying preferences, whereas people either have a preference a for both strategies intuitive and deliberate decision-making (PID-S plus) or without a preference for any strategy (PID-S minus). The PID-D group was most represented (N = 22), followed by PID-S minus (N = 17). There are no significant differences in the incidence of PID scores by age group (Pearson chi-square = 6.495; DF = 9; *p-value* = 0.690).

Table 58. *Crosstab PID Score * Age Group*

		Age Group				Sum
		up to 30 Y.	up to 40 Y.	up to 50 Y.	over 50 Y.	
PID Score	PID-D	6	8	5	3	22
	PID-I	0	6	4	1	11
	PID-S minus	2	7	5	3	17
	PID-S plus	2	4	3	0	9
Sum		10	25	17	7	59

Following A4 it would be an expected that rational people would end up with the highest number of correct answers. However, the data gathered in the experiment indicated that intuitive people might have a slight advantage in terms of predictive quality.

Table 59. *Comparison of All Predictions Grouped by PID Scale Score*

	Participants	Correct	Wrong	Sum	Percentage of correct answers
PID-D	22	2610	2514	5124	50.9%
PID-I	11	1320	1245	2565	51.5%
PID-S minus	17	2098	1967	4065	51.6%
PID-S plus	9	1082	883	1965	55.1%

The overall accuracy of people with a preference for deliberative decision-making was 50.9% and the overall accuracy of people with a preference for intuitive decision-making was 51.5% (see Table 59). Obviously this is not a very significant difference. The direct comparison of PID-D and PID-I prediction quality tested using a Chi-Square test results in a Chi-square: 0.189 and *p-value=0.66.* A comparison that includes all PID-types and in particular the result that a preference a for both strategies intuitive and deliberate decision-making (PID-S plus) led to significant better predictions, might be more interesting. This direct comparison of all four categories' predictive quality tested with a Chi-Square test results in a Chi-square: 10.084 and *p-value=0.018.* These results are to some extent in agreement with the explanatory scheme of Philip Tetlock and his interpretation of the metaphor of *The Hedgehog and the Fox.* In his analysis the aggregated success rate of the Foxes' predictions was considerably better compared with Hedgehogs and “Foxes were not especially likely to endorse particular substantive positions on rationality, level of analysis, macroeconomics, or foreign policy” (Tetlock, 2005, p. 106). Apparently, it is an advantage for forecasters to apply multiple strategies.

In summary it can be concluded that A4 is not supported by the data gathered in the experiment. However, an inverted version of the assumption: Intuitive people are better at financial decision-making compared with rational people would be supported by the data, although the data is not significant in a direct comparison of PID-I and PID-D results at a level of significance of 0.05. Higher significance could be observed in the direct comparison of all four categories. Predictions of PID-S plus participants are apparently of significantly higher accuracy.

Area of Discussion A5: Educational level has a major impact on the ability to predict stock prices.

Education is a very important component that influences people's decision-making strategies and it might be a sensible assumption that the level of education has an influence on decision-making ability and predictive accuracy (see also 2.3 Decision-Making and Forecasting). Accordingly, A5 assumes that the educational level of the participants has a major impact on the ability to predict stock prices. The data gathered in the experiment did —at least to some extent—conform with A5. Even so, there was only a weakly significant difference in the predictive accuracy of people with and without university degrees (Chi-square 2.967, *p-value*=0.084). People without university degrees were, with 50.4% correct answers, even slightly less often correct than people with university degrees, who produced just 52.1% correct answers (see Table 60).

Table 60. *Comparison of Predictions Grouped by Level of Education*

Level of Education	Participants	Correct	Wrong	Sum	Percentage of correct answers
without University Degree	18	2026	1994	4020	50.4%
with University Degree	31	3774	3471	7245	52.1%

A closer look at predictive accuracy according to educational level among the lay group of the EDG and NFG group in particular, with comparable group size and education level, exhibits an even more erratic result.

Table 61. *Predictions of Lay People Grouped by Level of Education and Group Design*

Level of Education	Group	Participants	Correct	Wrong	Sum	Percentage of correct answers
without University Degree	EDG	8	1007	898	1905	52.9%
with University Degree	EDG	13	1619	1561	3180	50.9%
without University Degree	IG	1	119	121	240	49.6%
with University Degree	IG	6	618	702	1320	46.8%
without University Degree	NFG	9	900	975	1875	48.0%
with University Degree	NFG	12	1537	1208	2745	56.0%

While it appears that participants with university degrees are not in all cases better, it has to be noted that statistically significant differences in the frequency of correct predictions could be observed for the 1-month period (*p-value<0.0001*), see the following crosstabs (Tables 62, 63, 64). Participants with university degrees were represented in the data set for the 1-month period significantly more in terms of correct forecasts than expected (assuming a random distribution). Altogether, at least some indication of a causal relationship between level of education and prediction is supported by the data from the experiment. Thus A5 is also supported by the data from the experiment conducted and should be accepted.

Table 62. *Crosstab University Degree * 1-Week Predictions*

Crosstab			1-Week Predictions		
			wrong	correct	Sum
University Degree	No	Frequency	703a	682a	1385
		Expected Frequency	693.3	691.7	1385.0
		% in 1-Week Predictions	30.7%	29.9%	30.3%
	Yes	Frequency	1586a	1602a	3188
		Expected Frequency	1595.7	1592.3	3188.0
		% in 1-Week Predictions	69.3%	70.1%	69.7%
Sum		Frequency	2289	2284	4573
		Expected Frequency	2289.0	2284.0	4573.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	.393[a]	1	.531
Frequency of valid Data Sets	4573		

			University Degree		
			No	Yes	Sum
1-Week Predictions	wrong	Frequency	703	1586	2289
		Expected Frequency	693.3	1595.7	2289.0
		% with University Degree	50.8%	49.7%	50.1%
	correct	Frequency	682	1602	2284
		Expected Frequency	691.7	1592.3	2284.0
		% with University Degree	49.2%	50.3%	49.9%
Sum		Frequency	1385	3188	4573
		Expected Frequency	1385.0	3188.0	4573.0
		% with University Degree	100.0%	100.0%	100.0%

Table 63. *Crosstab University Degree * 1 Month Predictions*

Crosstab			1-Month Predictions		
			wrong	correct	Sum
University Degree	No	Frequency	760[a]	625[b]	1385
		Expected Frequency	693.9	691.1	1385.0
		% in 1-Month Predictions	33.2%	27.4%	30.3%
	Yes	Frequency	1531[a]	1657[b]	3188
		Expected Frequency	1597.1	1590.9	3188.0
		% in 1-Month Predictions	66.8%	72.6%	69.7%
Sum		Frequency	2291	2282	4573
		Expected Frequency	2291.0	2282.0	4573.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	18.121[a]	1	.000
Frequency of valid Data Sets	4573		

			University Degree		
			No	Yes	Sum
1-Month Predictions	wrong	Frequency	760	1531	2291
		Expected Frequency	693.9	1597.1	2291.0
		% with University Degree	54.9%	48.0%	50.1%
	correct	Frequency	625	1657	2282
		Expected Frequency	691.1	1590.9	2282.0
		% with University Degree	45.1%	52.0%	49.9%
Sum		Frequency	1385	3188	4573
		Expected Frequency	1385.0	3188.0	4573.0
		% with University Degree	100.0%	100.0%	100.0%

Table 64. *Crosstab University Degree * 3 Month Predictions*

Crosstab			3-Month Predictions: wrong	3-Month Predictions: correct	Sum
University Degree	No	Frequency	600[a]	785[a]	1385
		Expected Frequency	614.5	770.5	1385.0
		% in 3-Month Predictions	29.6%	30.9%	30.3%
	Yes	Frequency	1429[a]	1759[a]	3188
		Expected Frequency	1414.5	1773.5	3188.0
		% in 3-Month Predictions	70.4%	69.1%	69.7%
Sum		Frequency	2029	2544	4573
		Expected Frequency	2029.0	2544.0	4573.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	.884[a]	1	.347
Frequency of valid Data Sets	4573		

			University Degree: No	University Degree: Yes	Sum
3-Month Predictions	wrong	Frequency	600	1429	2029
		Expected Frequency	614.5	1414.5	2029.0
		% with University Degree	43.3%	44.8%	44.4%
	correct	Frequency	785	1759	2544
		Expected Frequency	770.5	1773.5	2544.0
		% with University Degree	56.7%	55.2%	55.6%
Sum		Frequency	1385	3188	4573
		Expected Frequency	1385.0	3188.0	4573.0
		% with University Degree	100.0%	100.0%	100.0%

Area of Discussion A6: Women are better at assessing market sentiment and outperform with short term predictions.

Following the stereotype that female persons are better at evaluating emotions in people and society, it might be logically consistent to assume that they are also better at assessing market sentiment. Accordingly, they should be able to provide higher predictive quality for short term predictions than male participants. Since all expert participants in this experiment are male, an analysis of the lay people only is sensible in assessing A6. A direct comparison of the one-week predictions by female and male participants in the experiment shows that female participants were indeed able to provide predictions with slightly higher accuracy (see Table 65). Nevertheless, this data does not reach the level of statistical significance (Chi-square 3.1, *p-value*=0.078). Additional tests and more data might be helpful to clarify the relationship and effect. An extension of the testing of this assumption from just short term predictions (one-week) to the accuracy of the one-month and three-month predictions actually produces more significant results.

Table 65. *Comparison of 1 Week Predictions Grouped by Gender*

Gender	Participants	Correct	Wrong	Sum	Percentage of correct answers
Female Participants	21	861	819	1680	51.3%
Male Participants	28	1012	1063	2075	48.8%

A comparison of one-month predictions grouped by gender also reveals superior performance by female participants (see Table 66). For one-month predictions the significance test results in a Chi-square: 3.946, *p-value*=0.0469. Finally a comparison of the three-month predictions grouped by gender reveals an even higher level of superior performance by female participants (see Table 67). For three-month predictions the significance test results in a Chi-share: 9.979, *p-value*=0.0016.

Table 66. *Comparison of 1 Month Predictions Grouped by Gender*

Gender	Participants	Correct	Wrong	Sum	Percentage of correct answers
Female Participants	21	849	831	1680	50.5%
Male Participants	28	981	1094	2075	47.3%

Table 67. *Comparison of 3 Month Predictions Grouped by Gender*

Gender	Participants	Correct	Wrong	Sum	Percentage of correct answers
Female Participants	21	986	694	1680	58.7%
Male Participants	28	1111	964	2075	53.5%

A chi-square test of the aggregated one-week, one-month, and three-month predictions suggests a highly significant correlation with a chi-square of 14.681 and *p-value*<0.001 (see also Table 68). Apparently, female participants were significantly more accurate with their predictions in the experiment. It is also interesting that there were also considerably more female participants in the PID-S plus group.

Table 68. *Comparison of All Lay People Predictions Grouped by Gender*

Gender	Participants	Correct	Wrong	Sum	Percentage of correct answers
Female Participants	21	2696	2344	5040	53.5%
Male Participants	28	3104	3121	6225	49.9%

Table 69 shows a comparison of all lay participants grouped by gender and PID-Score. Within the participants it is noteworthy that there are relatively more female participants in the PID-I and PID-S plus categories. Male participants dominate the PID-S minus group, while the PID-D is relatively even distributed.

Table 69. *Comparison of All Lay People Predictions Grouped by Gender and PID-Score*

Gender	PID-Score	Partici-pants	Correct	Wrong	Sum	Percentage of correct answers	Relative proportion
Female Participants	PID-D	7	903	762	1665	54.2%	16.33%
Male Participants	PID-D	9	869	1021	1890	46.0%	15.75%
Female Participants	PID-I	5	700	605	1305	53.6%	11.66%
Male Participants	PID-I	5	506	559	1065	47.5%	8.75%
Female Participants	PID-S minus	4	463	482	945	49.0%	9.33%
Male Participants	PID-S minus	11	1406	1309	2715	51.8%	19.25%
Female Participants	PID-S plus	5	667	518	1185	56.3%	11.66%
Male Participants	PID-S plus	3	323	232	555	58.2%	5.25%

Area of Discussion A7: Financial analysts are consistently overoptimistic in their forecasts of covered shares.

The existing literature suggests that financial analysts are subject to a wide range of influences on their decision-making. In particular, overoptimistic estimates and preconditions are often laid to the charge of equity analysts (e.g., Bradshaw, Richardson, & Sloan, 2006; Cornett, Mehran, & Tehranian, 1998; Goedhart, Raj, & Saxena, 2010; Malmendier & Shanthikumar, 2007). Hence, an assumption (A7) to be tested is whether financial analysts are consistently overoptimistic in their forecasts of covered shares.

Since the data generated in the e-Delphi experiment provides insight from estimates and predictions by equity analysts and lay people, a direct comparison is possible. As a second step the data from the experiments also allows a comparison of estimates and predictions for companies within their coverage compared with estimates and predictions from their peers without active coverage of the respective company.

The data from the experiment does not support the assumption A7. Generally, lay people provided more optimistic predictions during the experiment. Optimistic predictions (stock price up) from analysts were, for one week, 43.3%, for one month 27.0%, and for three months 26.2%, while the percentage of optimistic predictions by lay participants was, for one week, 46.4%, for one month 44.2%, and for three months 54.2%.

Financial analysts were even less optimistic in their predictions. Only for five of fifteen possible shares/period combinations did the analysts with coverage provide more optimistic predictions (see Table 70). For six share/period combinations the analysts did not provide a single positive prediction during the whole experiment.

Table 70. *Predictions from Analysts' Coverage and the Average of All Participating Analysts*

Adidas	1-Week Predictions	1-Month Predictions	3-Month Predictions
All Analysts	59.8%	43.5%	43.4%
Analysts with coverage	50.0%	50.0%	75.0%
HeidelbergCement			
All Analysts	57.3%	36.7%	32.7%
Analysts with coverage	100.0%	100.0%	100.0%
RWE			
All Analysts	20.69%	4.39%	2.00%
Analysts with coverage	7.1%	0.0%	0.0%
Siemens			
All Analysts	48.2%	29.6%	33.2%
Analysts with coverage	26.3%	0.0%	0.0%
ThyssenKrupp			
All Analysts	22.5%	15.3%	15.3%
Analysts with coverage	16.7%	0.0%	0.0%

Area of Discussion A8: Life experience and age have an influence on stock price predictions. Older people are more risk averse.

Life experience might have an influence on the ability to appraise situations and also on the ability to predict future movements of stock prices (see also 2.3 Decision-Making and Forecasting). In this experiment, a proxy for life experience is the age of the participants. There was a tendency for older participants, i.e. the group over 50 years, to assess their own skills as lower than the average (see also figure 17). This might be an indication that older participants might indeed in a way be more risk averse than younger participants. However, there might also be other explanations for this lower self-assessment. During the interviews some of the older participants explained that they follow the news and stock market reports quite frequently, but still feel that they are not experts. One of the older participants explained quite emphatically that he is just interested in an appropriate dividend and long term growth.

Table 71. *Variables PID-D to PA3M according to Age Groups*

Age Group		PID-D	PID-I	Skill (Self-Assess.)	PA ALL	PA 1W	PA 1M	PA 3M
up to 30	Average	3.6222	2.5889	3.75	.499130	.4398	.4795	.5780
	H	10.0000	10.0000	10	10	10	10	10
	Standard deviation	1.3900	.9687	1.458	.0719623	.08352	.08418	.14033
	Minimum	.0000	.0000	2	.3750	.30	.32	.38
	Maximum	4.6667	3.3333	7	.6316	.57	.62	.74
	Median	4.0000	2.8333	3.25	.492600	.4472	.4758	.6014
up to 40	Average	3.6089	3.0267	5.48	.508572	.4984	.5033	.5241
	H	25.0000	25.0000	25	25	25	25	25
	Standard deviation	.7336	.5115	2.683	.0784637	.07683	.10127	.13719
	Minimum	1.5556	2.0000	1	.3767	.32	.34	.27
	Maximum	4.6667	3.8889	10	.6982	.62	.72	.76
	Median	3.6667	3.0000	5.50	.504800	.5100	.4706	.5700
up	Average	3.7516	3.1176	4.84	.527782	.5089	.5219	.5526

	H	17.0000	17.0000	16	17	17	17	17
	Standard deviation	.5336	.4738	2.047	.0757533	.07080	.11190	.11968
	Minimum	2.6667	2.3333	1	.3614	.38	.30	.29
to	Maximum	4.7778	4.0000	9	.6500	.67	.68	.74
50	Median	3.6667	3.1111	4.75	.523800	.4900	.5143	.5875
	Average	3.6825	2.9841	2.83	.521300	.5078	.4763	.5798
	H	7.0000	7.0000	6	7	7	7	7
	Standard deviation	.6618	.4689	1.571	.0443356	.04398	.09127	.06784
	Minimum	2.7778	2.2222	2	.4733	.45	.37	.49
ove	Maximum	4.5556	3.7778	6	.6000	.56	.65	.66
r 50	Median	3.5556	3.0000	2.25	.516700	.5000	.4750	.5500
	Average	3.6610	2.9736	4.72	.514017	.4926	.5014	.5480
	H	59.0000	59.0000	57	59	59	59	59
	Standard deviation	.8064	.6094	2.362	.0724599	.07555	.09971	.12570
	Minimum	.0000	.0000	1	.3614	.30	.30	.27
Su	Maximum	4.7778	4.0000	10	.6982	.67	.72	.76
m	Median	3.7778	3.0000	4.00	.506700	.5000	.4900	.5700

Table 71 provides a comprehensive overview of the different variables, in particular predictive accuracies for different periods. The data from the experiment does indeed indicate a positive correlation between age and amount of correct stock price predictions (see Figure 22).

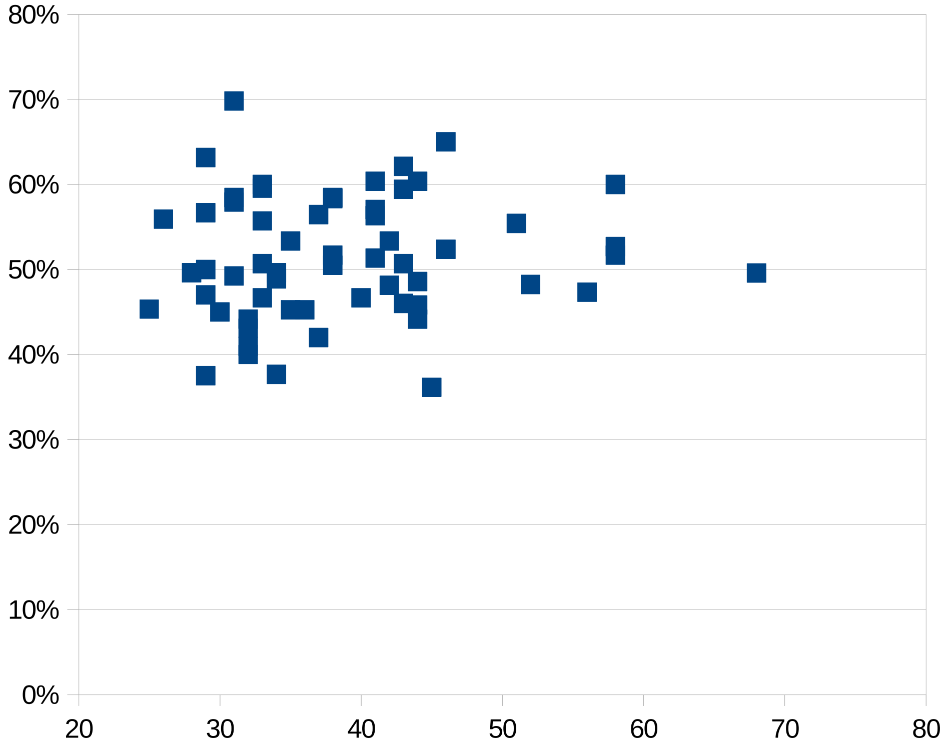


Figure 22: Overview of correct answers in percent and age of the participants

At first glance the data supports the assumption insofar as life experience and age have an influence on stock price predictions. However, it could be observed that there was a huge spread in predictive accuracy and some of the most accurate participants were comparably young (see Table 72). The participant with the best performance was 31 years old and the participant with the third best result only 29 years old, but some of the most inaccurate predictions also were also produced by rather young participants.

Table 72. *Overview Predictive Accuracy of the Ten Most and Least Accurate Participants*

Age of the participant	**Correct answers in percent**
31	69.82%
46	65.00%
29	63.16%
43	62.11%
41	60.37%
44	60.37%
33	60.00%
58	60.00%
33	59.56%
43	59.44%
…	…
30	45.00%
32	44.17%
44	44.17%
32	43.08%
37	42.00%
32	40.95%
32	40.00%
34	37.67%
29	37.50%
45	36.14%

A comparison of age cluster accuracy shows that participants younger than 30 years old had an average predictive accuracy of 50.66% (N=8), participants of 30 and younger than 40 had an average predictive accuracy of 57.53% (N=26), participants of 40 and younger than 50 had an average predictive accuracy of 56.4% (N=18), and participants 50 years and older had an average predictive accuracy of 52.13% (N=7). Obviously the data available to test A8 is not very extensive and further analysis with a bigger sample size could be useful. Nevertheless, might be worth making a note of the provisional finding that the impact of life experience, if there is any, is seemingly not very strong.

Still, there are significant differences in the frequencies of correct predictions for the 3 periods (1 week *p-value*=0.024, 1 month *p-value*=0.016, 3 month *p-value*<0.0001). The frequency of correct predictions is, in the age group up to 30 years for the 1-week periods, rarer than expected (assuming a random distribution). The percentage differences between wrong and correct prediction frequencies are significant (see Table 73). For the other age groups, however, no significant differences were observed.

Table 73. *Crosstab Age Groups * 1 Week Predictions*

Crosstab			1-Week Predictions wrong	correct	Sum
Age Group	up to 30 Years	Frequency	394$_a$	321$_b$	715
		Expected Frequency	357.9	357.1	715.0
		% in 1-Week Predictions	17.2%	14.1%	15.6%
	up to 40 Years	Frequency	965$_a$	973$_a$	1938
		Expected Frequency	970.1	967.9	1938.0
		% in 1-Week Predictions	42.2%	42.6%	42.4%
	up to 50 Years	Frequency	654$_a$	701$_a$	1355
		Expected Frequency	678.2	676.8	1355.0
		% in 1-Week Predictions	28.6%	30.7%	29.6%
	over 50 Years	Frequency	276$_a$	289$_a$	565
		Expected Frequency	282.8	282.2	565.0
		% in 1-Week Predictions	12.1%	12.7%	12.4%
Sum		Frequency	2289	2284	4573
		Expected Frequency	2289.0	2284.0	4573.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	9.410[a]	3	.024
Frequency of valid Data Sets	4573		

Considering the wrong and correct predictions per age group (see Table 74), it becomes apparent that only people up to 30 years had less than 50% correct predictions for a 1-week period (only 44.9%), but that there are slightly more than 50% correct predictions in the older age groups (between 50.2% and 51.7%).

Table 74. *Crosstab 1 Week Predictions * Age Groups*

			Age Group				
			up to 30 Years	up to 40 Years	up to 50 Years	over 50 Years	Sum
1-Week Predictions	wrong	Frequency	394	965	654	276	2289
		Expected Frequency	357.9	970.1	678.2	282.8	2289.0
		% in Age Group	55.1%	49.8%	48.3%	48.8%	50.1%
	correct	Frequency	321	973	701	289	2284
		Expected Frequency	357.1	967.9	676.8	282.2	2284.0
		% in Age Group	44.9%	50.2%	51.7%	51.2%	49.9%
Sum		Frequency	715	1938	1355	565	4573
		Expected Frequency	715.0	1938.0	1355.0	565.0	4573.0
		% in Age Group	100.0%	100.0%	100.0%	100.0%	100.0%

Participants with correct predictions in the age group over 50 years were, for the 1-month periods rarer than expected (assuming random distribution). The percentage differences between wrong and correct prediction frequencies are significant (see Table 75). For the other age groups, however, there were no significant differences.

Table 75. *Crosstab Age Groups * 1 Month Predictions*

			1-Month Predictions		
Crosstab			wrong	correct	Sum
Age Group	up to 30 Years	Frequency	377_a	338_a	715
		Expected Frequency	358.2	356.8	715.0
		% in 1-Month Predictions	16.5%	14.8%	15.6%
	up to 40 Years	Frequency	955_a	983_a	1938
		Expected Frequency	970.9	967.1	1938.0
		% in 1-Month Predictions	41.7%	43.1%	42.4%
	up to 50 Years	Frequency	649_a	706_a	1355
		Expected Frequency	678.8	676.2	1355.0
		% in 1-Month Predictions	28.3%	30.9%	29.6%
	over 50 Years	Frequency	310_a	255_b	565
		Expected Frequency	283.1	281.9	565.0
		% in 1-Month Predictions	13.5%	11.2%	12.4%
Sum		Frequency	2291	2282	4573

	Expected Frequency	2291.0	2282.0	4573.0
	% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	10.266[a]	3	.016
Frequency of valid Data Sets	4573		

A closer look at the wrong and correct predictions per age group (see Table 75) reveals that participants up to 30 years and over 50 years had less than 50% correct predictions for the one-month periods (47.3% respectively 45.1%). However, the other age groups were slightly above 50% correct predictions (between 50.7% and 52.1%).

Table 76. *Crosstab 1 Month Predictions * Age Groups*

			Age Group				
			up to 30 Years	up to 40 Years	up to 50 Years	over 50 Years	Sum
1-Week Predictions	wrong	Frequency	394	965	654	276	2289
		Expected Frequency	357.9	970.1	678.2	282.8	2289.0
		% in Age Group	55.1%	49.8%	48.3%	48.8%	50.1%
	correct	Frequency	321	973	701	289	2284
		Expected Frequency	357.1	967.9	676.8	282.2	2284.0
		% in Age Group	44.9%	50.2%	51.7%	51.2%	49.9%
Sum		Frequency	715	1938	1355	565	4573
		Expected Frequency	715.0	1938.0	1355.0	565.0	4573.0
		% in Age Group	100.0%	100.0%	100.0%	100.0%	100.0%

The frequency of correct predictions is, in the age group up to 30 years for the 3-month period, higher than expected (assuming a random distribution), the percentage differences between false and correct prediction frequencies are significantly different (see Table 77). In the group up to 40 years the percentages between false and correct predictions are also significantly different; however, the number of correct predictions for the 3 month periods in this age group were less frequent than expected. For the other age groups, however, there were no significant differences.

Table 77. *Crosstab Age Groups * 3 Month Predictions*

Crosstab			3-Month Predictions: wrong	3-Month Predictions: correct	Sum
Age Group	up to 30 Years	Frequency	268[a]	447[b]	715
		Expected Frequency	317.2	397.8	715.0
		% in 3-Month Predictions	13.2%	17.6%	15.6%
	up to 40 Years	Frequency	916[a]	1022[b]	1938
		Expected Frequency	859.9	1078.1	1938.0
		% in 3-Month Predictions	45.1%	40.2%	42.4%
	up to 50 Years	Frequency	602[a]	753[a]	1355
		Expected Frequency	601.2	753.8	1355.0
		% in 3-Month Predictions	29.7%	29.6%	29.6%
	over 50 Years	Frequency	243[a]	322[a]	565
		Expected Frequency	250.7	314.3	565.0
		% in 3-Month Predictions	12.0%	12.7%	12.4%
Sum		Frequency	2029	2544	4573
		Expected Frequency	2029.0	2544.0	4573.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	20.749[a]	3	.000
Frequency of valid Data Sets	4573		

Considering the wrong and correct predictions per age group (see Table 78), it becomes evident that all ages achieved more than 50% correct predictions for the 3 month periods (between 62.5% and 52.7%). Overall, 55.6% correct predictions were made (all age groups). The predictive accuracy of the age group up to 30 years and the group over 50 years was above this value.

Table 78. *Crosstab 3 Month Predictions * Age Groups*

			Age Group				Sum
			up to 30 Years	up to 40 Years	up to 50 Years	over 50 Years	
3-Month Predictions	wrong	Frequency	268	916	602	243	2029
		Expected Frequency	317.2	859.9	601.2	250.7	2029.0
		% in Age Group	37.5%	47.3%	44.4%	43.0%	44.4%
	correct	Frequency	447	1022	753	322	2544
		Expected Frequency	397.8	1078.1	753.8	314.3	2544.0
		% in Age Group	62.5%	52.7%	55.6%	57.0%	55.6%
Sum		Frequency	715	1938	1355	565	4573
		Expected Frequency	715.0	1938.0	1355.0	565.0	4573.0
		% in Age Group	100.0%	100.0%	100.0%	100.0%	100.0%

Area of Discussion A9: Analysts are better than lay people in bull markets, but lose that advantage in bear markets.

There is some discussion in academia and among practitioners whether analysts are over-optimistic (Cornett et al., 1998; Goedhart et al., 2010; Kellerman, 2014; Michel & Pandes, 2012). Goedhart et al. argue that “analysts have been persistently over-optimistic for the past 25 years” (2010, p. 16). These statements usually compare the analysts' estimates with actual conditions. Still it remains unclear how to improve their estimates. This might lead to the question whether analysts are better than lay people in bull markets, but lose that advantage in bear markets. The data from the experiment allows a direct comparison of the predictions by financial analysts and lay people. An analysis of this data might also contribute to the discussion about analysts' optimism. During the period of the experiment there were 100 different measurements for predictions for each period (1-week, 1-month, and 3-month). It turned out that of the 100 measurements for 1-week predictions 52 stock price movement were upwards, and in 48 periods prices were falling. Of the 100 measurements for 1-month predictions 59 stock price movements were upwards, and in 41 periods prices were falling. Of the 100 measurements for 3-month predictions 64 stock price movement here upwards and in 36 periods prices were falling. A differentiated analysis of the data sets with falling and rising prices allows assessing the prediction quality in bull and bear markets.

Table 79. *Comparison of Group Results in Periods with Prices on the Rise*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Right	19	14	16	32	17	25
	Wrong	21	35	34	19	21	23
	Excluded	12	3	2	1	14	4
	Right	47.5%	28.6%	32.0%	62.7%	44.7%	52.1%
	Wrong	52.5%	71.4%	68.0%	37.3%	55.3%	47.9%
1-Month	Right	7	20	20	27	32	26
	Wrong	40	37	34	30	19	31

	Excluded	12	2	5	2	8	2
	Right	14.9%	35.1%	37.0%	47.4%	62.7%	45.6%
	Wrong	85.1%	64.9%	63.0%	52.6%	37.3%	54.4%
3-Month	Right	7	51	35	36	51	32
	Wrong	45	12	26	25	4	30
	Excluded	12	1	3	3	9	2
	Right	13.5%	81.0%	57.4%	59.0%	92.7%	51.6%
	Wrong	86.5%	19.0%	42.6%	41.0%	7.3%	48.4%
All Periods	Right	33	85	71	95	100	83
	Wrong	106	84	94	74	44	84
	Excluded	36	6	10	6	31	8
	Right	23.7%	50.3%	43.0%	56.2%	69.4%	49.7%
	Wrong	76.3%	49.7%	57.0%	43.8%	30.6%	50.3%

The difference between the predictive accuracy for the periods of rising prices is significant (Chi-square: 68.582; DF=5; *p-value*=0). The analysis revealed that for the periods of rising prices the accuracy of the financial analysts was, with only 23.7% correct predictions, worst of all groups, while financial professionals, with a predictive accuracy of 69.4% performed best (see Table 79). The difference in predictive accuracy for the lay groups is only weakly significant (Chi-square: 5.819; DF=2; *p-value*=0.055).

Table 80. *Comparison of Group Results in Periods with Falling Prices*

		AG	EDG	IG	NFG	PG	Expert
1-Week	Right	29	28	25	25	13	29
	Wrong	12	19	14	19	31	13
	Excluded	7	1	9	4	4	6
	Right	70.7%	59.6%	64.1%	56.8%	29.5%	52.1%
	Wrong	29.3%	40.4%	35.9%	43.2%	70.5%	47.9%
1-Month	Right	38	22	15	22	24	31
	Wrong	3	16	16	15	8	2
	Excluded	0	3	10	4	9	8
	Right	92.7%	57.9%	48.4%	59.5%	75.0%	93.9%
	Wrong	7.3%	42.1%	51.6%	40.5%	25.0%	6.1%

3-Month	Right	34	19	7	23	11	28
	Wrong	1	15	25	12	18	0
	Excluded	1	2	4	1	7	8
	Right	97.1%	55.9%	21.9%	65.7%	37.9%	100.0%
	Wrong	2.9%	44.1%	78.1%	34.3%	62.1%	0.0%
All Periods	Right	101	69	47	70	48	88
	Wrong	16	50	55	46	57	15
	Excluded	8	6	23	9	20	22
	Right	86.3%	58.0%	46.1%	60.3%	45.7%	85.4%
	Wrong	13.7%	42.0%	53.9%	39.7%	54.3%	14.6%

The difference in predictive accuracy for the periods of rising prices is significant (Chi-square: 77.753; DF=5; *p-value*=0). The analysis revealed that for the periods of falling prices the accuracy of the financial analysts was, with 86.3% correct predictions, best of all groups, while financial professionals, with a predictive accuracy of 45.7% performed worst of all groups (see Table 80). The difference in predictive accuracy for the lay groups is only weakly significant (Chi-square: 5.046; DF=2; *p-value*=0.080).

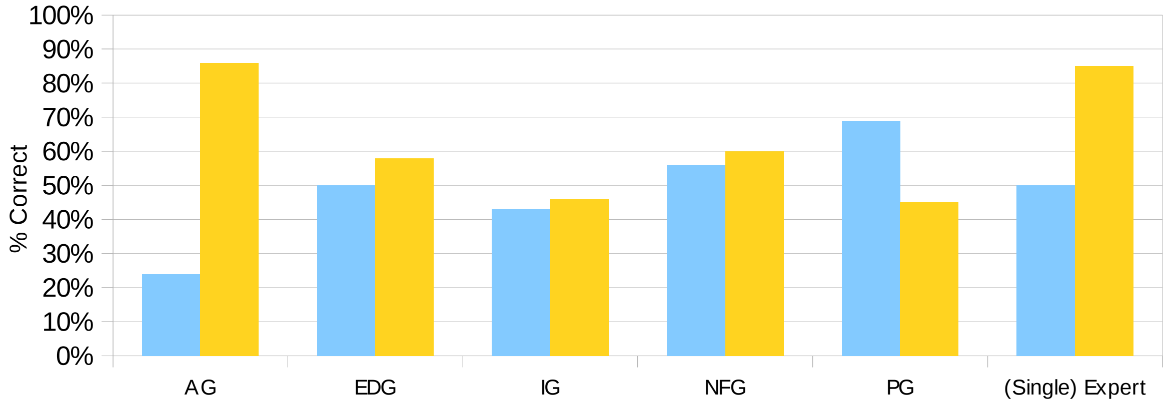


Figure 23: Predictive accuracy in bull (blue) and bear (yellow) markets.

The direct comparison of the predictions for periods of falling and rising prices shows that analysts had the greatest difference in predictive accuracy (see Figure 23). All groups except the financial professionals had higher predictive accuracy in periods of falling prices. An interesting observation might be that the predictive accuracy of financial analysts and financial professionals was oppositional in periods of rising and falling share prices. Notwithstanding, assumption A9 is not supported by the data of the experiment.

Area of Discussion A10: People who are interested in the stock market are able to provide better predictions.

The data from the experiment doesn't support the assumption (A10) that people who are interested in the stock market are able to provide better predictions. Participants who stated in the interviews that they had no or only very little interest in stock markets had an average predictive accuracy of 51.5% (*N*=24), while participants who stated that they were interested in stock markets had an average predictive accuracy of 51.1% (*N*=29). Six participants did not answer the question: their average accuracy was 52.3%.

All the experts stated that they are interested in stock markets. An analysis excluding the experts shows that lay participants who stated in the interviews that they have no or only very little interest in stock markets had an average predictive accuracy of 51.5% (*N*=24), while participants who stated that they are interested in stock markets had an average predictive accuracy of 50.0% (*N*=19).

This results suggest the conclusion that whether people are interested or not/very little interested in the stock market has no impact on their ability to predict stock price movements. This finding is very interesting in light of the idea that a driver for swarm intelligence is diversity in the group design (see also2.3 Decision-Making and Forecasting and 2.4 Decision Support Systems (DSS)). While self-selection might be an important factor in participation in Internet groups, there might be also a tendency for like-minded people to meet in these groups and create, instead of an exchange of different opinions, just a reconfirmation of existing prejudice. In certain situations it might be beneficial to the overall group decision-making to have people with different mind-sets.

Area of Discussion A11: People who think that they know more about the stock market are able to make better forecasts.

The data from the experiment does not support the assumption (A11) that people who think that they know more about the stock market are able to make better forecasts (see also 2.2 Equity Research). An analysis of the data (see also Figure 24) even showed a slightly negative correlation of -0,006. Apparently one's own assessment of stock market expertise does not impact the ability to predict stock prices.

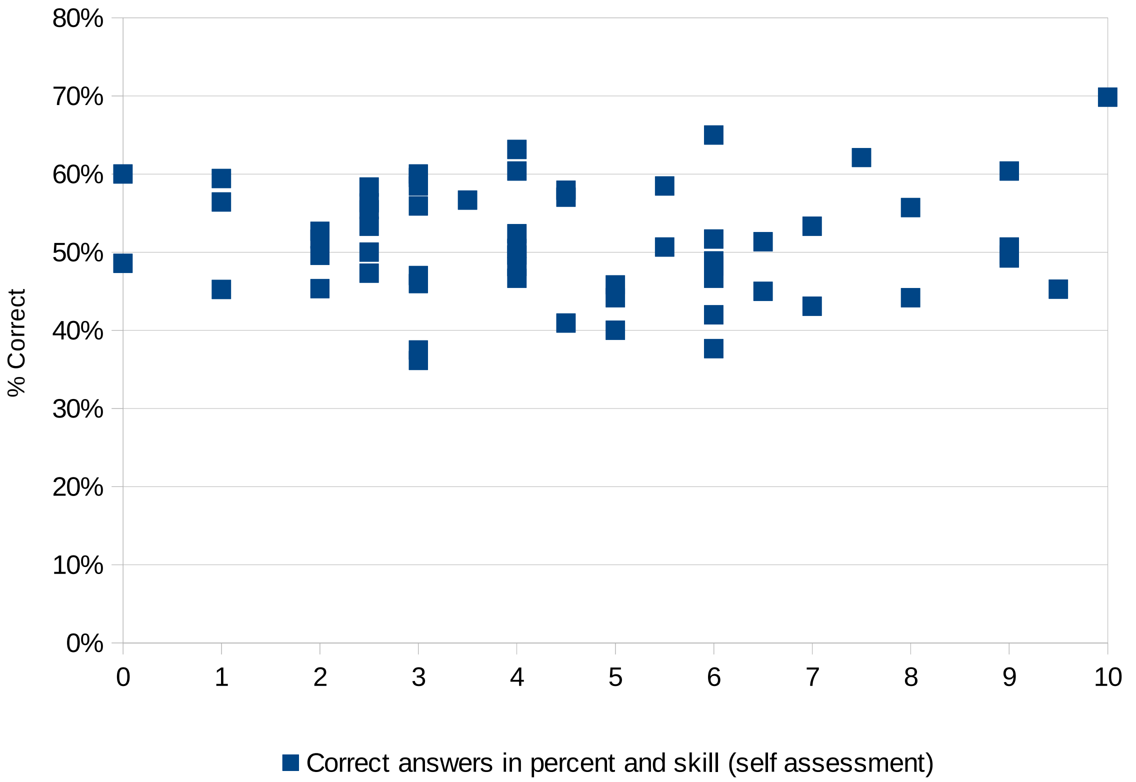


Figure 24: Overview of correct answers in percent and skill (self assessment)

Area of Discussion A12: Predictions with a higher level of confidence are generally better than predictions with low confidence.

Among participants who were unsure or very confident in their decisions (according to self-evaluation) there were significant differences in the percentages of correct predictions (see same subscript letters in table 81). The data for the 1-week predictions showed that respondents who stated that they were confident ("sure") in their prognosis were represented at 19.8% with wrong and only about 17.5% with correct answers. Still, it is true that for the 1-week predictions no significant differences in the frequency of correct or incorrect predictions depended on confidence (Pearson chi-square = 3.119; FG = 3; *p-value = 0.374*), as in the other assessments (of "not sure at all" to "rather sure") no significant differences in the frequency of correct or incorrect predictions were discovered.

Table 81. *Crosstab Confidence Level and One Week Prediction Accuracy*

Crosstab Confidence Level* PA 1-Week			1-Week Correct		Sum
			wrong	correct	
Confidence Level	not sure at all	Frequency	291_a	332_a	623
		Expected Frequency	311.8	311.2	623.0
		% Correct	12.7%	14.5%	13.6%
	rather unsure	Frequency	575_a	604_a	1179
		Expected Frequency	590.1	588.9	1179.0
		% Correct	25.1%	26.4%	25.8%
	rather sure	Frequency	965_a	935_a	1900
		Expected Frequency	951.0	949.0	1900.0
		% Correct	42.2%	40.9%	41.5%
	sure	Frequency	454_a	399_b	853
		Expected Frequency	427.0	426.0	853.0
		% Correct	19.8%	17.5%	18.7%
	absolutely sure	Frequency	4_a	14_b	18
		Expected	9.0	9.0	18.0

		Frequency			
		% Correct	0.2%	0.6%	0.4%
Sum		Frequency	2289	2284	4573
		Expected Frequency	2289.0	2284.0	4573.0
		% Correct	100.0%	100.0%	100.0%
Each subscript letter indicates a subset of the 1-week predictive accuracy category, where columns do not differ significantly on .05-level of significance.					

The data for the 1-month predictions showed that respondents who stated that they were confident (coded as “sure”) in their prognosis were represented at 24.1% with wrong and about 27.4% with correct answers; in contrast, moderately confident (“rather sure”) participants had a share of 43.2% in wrong predictions and 39.9% correct predictions. For all other groups no significant differences in the frequency of correct or incorrect predictions were discovered (see table 82).

There are significant differences in the frequency of correct or incorrect predictions for the 1-month predictions depending on confidence (Pearson chi-square = 10.889; DF = 4; *p-value = 0.028*).

Table 82. *Crosstab Confidence Level and One Month Prediction Accuracy*

Crosstab Confidence Level* PA 1-Month			1-Month Correct		Sum
			wrong	correct	
Confidence Level	not sure at all	Frequency	299_a	324_a	623
		Expected Frequency	312.1	310.9	623.0
		% Correct	13.1%	14.2%	13.6%
	rather unsure	Frequency	553_a	626_b	1179
		Expected Frequency	590.7	588.3	1179.0
		% Correct	24.1%	27.4%	25.8%
	rather sure	Frequency	990_a	910_b	1900
		Expected Frequency	951.9	948.1	1900.0

		% Correct	43.2%	39.9%	41.5%
	sure	Frequency	442_a	411_a	853
		Expected Frequency	427.3	425.7	853.0
		% Correct	19.3%	18.0%	18.7%
	absolutely sure	Frequency	7_a	11_a	18
		Expected Frequency	9.0	9.0	18.0
		% Correct	0.3%	0.5%	0.4%
Sum		Frequency	2291	2282	4573
		Expected Frequency	2291.0	2282.0	4573.0
		% Correct	100.0%	100.0%	100.0%

Each subscript letter indicates a subset of the 1-month predictive accuracy category, were columns do not differ significantly on .05-level of significance.

The data for the 3-month predictions showed that there are no significant differences in the frequency of correct or incorrect predictions (see table 83).

There are <u>no</u> significant differences in the frequency of correct or incorrect predictions for the 1-month predictions depending on confidence (Pearson chi-square = 3.333; DF = 4; *p-value = 0.504*). Still, very confident participants (stating that they are "absolutely sure") had a slightly higher proportion of correct predictions.

Table 83. *Crosstab Confidence Level and Three Month Prediction Accuracy*

Crosstab Confidence Level* PA 3-Month			3-Month Correct wrong	correct	Sum
Confidence Level	not sure at all	Frequency	281_a	342_a	623
		Expected Frequency	276.4	346.6	623.0
		% Correct	13.8%	13.4%	13.6%
	rather unsure	Frequency	508_a	671_a	1179
		Expected Frequency	523.1	655.9	1179.0
		% Correct	25.0%	26.4%	25.8%

	rather sure	Frequency	868_a	1032_a	1900
		Expected Frequency	843.0	1057.0	1900.0
		% Correct	42.8%	40.6%	41.5%
	sure	Frequency	365_a	488_a	853
		Expected Frequency	378.5	474.5	853.0
		% Correct	18.0%	19.2%	18.7%
	absolutely sure	Frequency	7_a	11_a	18
		Expected Frequency	8.0	10.0	18.0
		% Correct	0.3%	0.4%	0.4%
Sum		Frequency	2029	2544	4573
		Expected Frequency	2029.0	2544.0	4573.0
		% Correct	100.0%	100.0%	100.0%

Each subscript letter indicates a subset of the one-month predictive accuracy category, where columns do not differ significantly on 0.05-level of significance.

Overall the data from the experiment does support the assumption (A12) that predictions with higher confidence are generally better than predictions with low confidence (see Figure 25). All participants were asked with every company stock price prediction: “How confident are you in this prediction?”. It was a mandatory question with a Likert-type scale from 1 (not sure at all) to 5 (absolutely sure). The data revealed that there is a positive correlation between the confidence of the predictions provided by the participants and the accuracy of the prediction. The correlation is 0.56 for the one-week predictions, but still clearly positive for the one-month predictions (0.39), and most strongly positive for the three-month predictions (0.75).

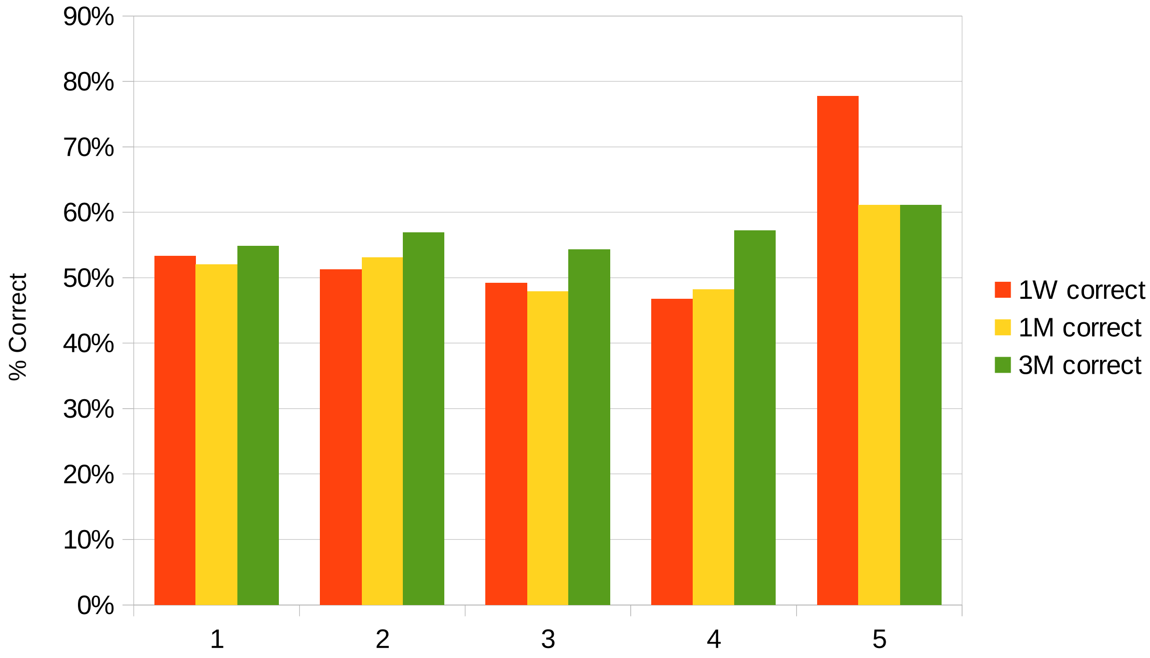


Figure 25: Percentage of correct answers grouped by confidence in the prediction given

An interesting observation is that the positive correlation is caused mainly by the predictions with the highest confidence. Apparently predictions with high confidence are the most accurate ones. Another observation was that people seem to be very restricted in their use of the highest confidence rating: only 54 of 4573 predictions (see also Table 84) provided were labelled with the highest confidence rating 5 (absolutely sure).

Table 84. *Predictive Accuracy According to Expressed Commitment*

Commit-ment	1-week correct	1-week wrong	1-week accuracy	1-month correct	1-month wrong	1-month accuracy	3-month correct	3-month wrong	3-month accuracy
1	332	291	53.3%	324	299	52.0%	342	281	54.9%
2	604	575	51.2%	626	553	53.1%	671	508	56.9%
3	935	965	49.2%	910	990	47.9%	1032	868	54.3%
4	399	454	46.8%	411	442	48.2%	488	365	57.2%
5	14	4	77.8%	11	7	61.1%	11	7	61.1%

On the basis of that information, an analysis of the data with confidence ratings from 1 to 4 only was conducted. The correlation for the confidence ratings 1 to 4, excluding 5, is strongly negative for the one-week predictions (-1.00), still clearly negative for the one-month predictions (-0.81), but positive for the three-month predictions (0.39). It seems as if participants were quite reluctant to opt for confidence level 5, but if so there was considerable superior performance, in particular for the short term one-week predictions (Chi-square=5.56 and *p-value*=0.018).

Area of Discussion A13: When people express a higher upside or downside (in terms of price to target price difference) the predictions are better.

In order to test assumption A13 the data gathered in the experiment was grouped into 8 categories according to the price target estimates provided by the participants. The analysis produced the following category groups: “<= -20%”, “> -20% and <= -10%”, “> -10% and <= -5%”, “> -5% and <= 0%”, “> 0% and <= 5%”, > 5% and <= 10%”, “> 10% and <= 20%” to “>20%”. Table 85 shows an overview of predictive accuracy grouped by price target estimates by the participants. Overall 4216 forecast data points could be included in the analysis. From 5900 possible forecasts in the experiment 1684 were missing or did not include a price target and are accordingly excluded in the analysis of A13.

Apparently the highest accuracy was for the 3-month predictions and a price target of more than 20 percent with an accuracy of 84.1%. This finding goes very well with the assumption that when people express a higher upside or downside (in terms of price to target price difference) the predictions are better. However, it is necessary to consider all the data to get a better picture. It is also true that the second highest result for the 3-month predictions was for a price target from 0% to 5% with an accuracy of 68.3%.

Table 85. *Prediction Accuracy According Grouped by 3 Month Price Targets*

	1-Week results			1-Month results			3-Month results			
Price target (3-month)	Correct	Wrong	Correct (%)	Correct	Wrong	Correct (%)	Correct	Wrong	Correct (%)	Sum
<= -20%	33	27	55.0%	37	23	61.7%	35	25	58.3%	60
> -20% and <= -10%	172	168	50.6%	151	189	44.4%	131	209	38.5%	340
> -10% and <= -5%	281	195	59.0%	241	235	50.6%	198	278	41.6%	476
> -5% and <= 0%	559	555	50.2%	524	590	47.0%	451	663	40.5%	1114
> 0% and <= 5%	840	951	46.9%	895	896	50.0%	1223	568	68.3%	1791

> 5% and <= 10%	139	145	48.9%	149	135	52.5%	192	92	67.6%	284
> 10% and <= 20%	37	45	45.1%	48	34	58.5%	50	32	61.0%	82
>20%	31	38	44.9%	43	26	62.3%	58	11	84.1%	69

Overall the data presents a mixed picture: while for the one-week predictions the negative price targets coincide with higher predictive accuracy, it is true that for the longer terms (1-month and 3-month) positive price targets coincided with higher predictive accuracy in the experiment. A more aggregated analysis of the results was carried out with the data grouped in only two categories: A category with moderate price targets (from -5% to 5%) and a category with more extreme price targets (less than -5% and more than 5%). Table 86 shows that for the one-week and one-month results the more extreme target prices were connected with more accurate predictions, but that on a 3-month basis the moderate price targets were more accurate.

Table 86. *Aggregated Predictive Accuracy According Grouped by 3 Month Price Targets*

	1-Week results			1-Month results			3-Month results			
Price target (3-month)	Correct	Wrong	Correct (%)	Correct	Wrong	Correct (%)	Correct	Wrong	Correct (%)	Sum
-5% to 5%	1399	1506	48.2%	1419	1486	48.85 %	1674	1231	57.6%	1311
less than -5% and more then 5%	693	618	52.9%	669	642	51.0%	664	647	50.6%	2905

Even though the data revealed a somewhat mixed picture it was still true that the most extreme price targets coincided with higher predictive accuracy, at least for the longer periods (see Table 87). While the one-week results have no statistical significance (Chi-square=0 and *p-value=1),* the one-month results are highly significant (Chi-square=7.811 and *p-value=0.005)* and the three-month results are even more significant (Chi-square=14.058 and *p-value<0.001)*.

Table 87. *Predictive Accuracy (with High 3 Month Price Targets)*

	1-Week results			1-Month results			3-Month results			
Price target (3-month)	Correct	Wrong	Correct (%)	Correct	Wrong	Correct (%)	Correct	Wrong	Correct (%)	Sum
Average	64	65	49.6%	80	49	62.0%	93	36	72.1%	129
less than -20% and more than 20%	2092	2124	49.6%	2088	2128	49.5%	2338	1878	55.5%	4216

Area of Discussion A14: Predictions based on fundamental (or technical?) analysis are superior to intuitive predictions.

All participants were asked to provide the basis of their decision and to mention influencing factors, i.e. basic decision-making principles and information sources. In the main experiment nine options were provided, which were based on the information gathered in the pilot run. Answers from the pilot run were clustered in nine groups (see Table 88) and presented as easy to understand options for the lay participants in the main experiment. Additionally, there was a free text entry box to provide the participants with an option to mention additional influencing factors in their individual decision-making. Multiple answers to this question were allowed (see Table 89). It was not mandatory for participants to tick a box or to enter a text, but most participants were willing to provide an explanation. For 4545 data sets (4573 data sets were actually submitted out of 5900 options to submit) participants chose at least one answer as rationale for their prediction.

Table 88. *Clusters of different decision-making fundamentals/influences*

Financial ratios	Market cap, P/E, dividend yields etc.	SQ001
Fundamental analysis	Discounted cash flow, dividend discount model, peer group analysis etc.	SQ002
Group results	Feedback from the e-Delphi group (last week or first round)	SQ003
Company	Products, brands, customers, innovations, company development	SQ004
Intuition	Like gut feeling, instinct, guess	SQ005
Market sentiment	General market situation and market outlook	SQ006
Experts	Financial analysts and other expert opinions	SQ007
News	Including daily press, Internet, business- and finance news	SQ008
Technical analysis	Index development, price-movement, momentum etc.	SQ009

More participants mentioned intuition as the basis for their decision than any other reason. Of the 4573 answers submitted, a total of 3844 (84.06%) were based, at least partly, on intuition i.e., reasons like gut feeling, instinct, guess. Only very few participants used the option to provide additional influencing factors (1.64%) and if so, it was often just a concretisation of other reasons like a particular newspaper, which could also have been put into category SQ008.

Participants with a definite basis for their decision had a higher likelihood of being correct with their predictions than decisions based on other approaches or information sources.

Table 89. *Number of correct predictions by decision-making principles*

1-Week	SQ001	SQ002	SQ003	SQ004	SQ005	SQ006	SQ007	SQ008	SQ009	other
Wrong	199	74	202	413	1935	1049	233	727	344	41
Correct	194	101	162	400	1909	978	272	682	314	34
1-Month	SQ001	SQ002	SQ003	SQ004	SQ005	SQ006	SQ007	SQ008	SQ009	other
Wrong	177	78	178	417	1937	1044	224	689	279	43
Correct	216	97	186	396	1907	983	281	720	379	32
3-Month	SQ001	SQ002	SQ003	SQ004	SQ005	SQ006	SQ007	SQ008	SQ009	other
Wrong	181	87	132	381	1744	955	194	626	269	29
Correct	212	88	232	432	2100	1072	311	783	389	46

For the one-week predictions the most accurate predictions included fundamental analysis (SQ002) as a basis for decision-making (see Table 90). One-week predictions based on that were significantly better than by chance (Chi-square: 4.166 and *p-value=0.041)*.

Table 90. C*orrect predictions by decision-making principles and information sources (in Percent)*

	SQ001	SQ002	SQ003	SQ004	SQ005	SQ006	SQ007	SQ008	SQ009	other
1-Week	49,4%	57,7%	44,5%	49,2%	49,7%	48,2%	53,9%	48,4%	47,7%	45.3%
1-Month	55,0%	55,4%	51,1%	48,7%	49,6%	48,5%	55,6%	51,1%	57,6%	42.7%
3-Month	53,9%	50,3%	63,7%	53,1%	54,6%	52,9%	61,6%	55,6%	59,1%	61.3%
Average	52,8%	54,5%	53,1%	50,3%	51,3%	49,9%	57,0%	51,7%	54,8%	49.8%

For the one-month predictions the most accurate predictions included technical analysis (SQ009) as the basis of the decision-making. One-month predictions based on that were significantly better than by chance (Chi-square: 15.198 and *p-value<0.0001)*.

However, other influencing factors such as expert opinions (SQ007) (Chi-square: 28.685 and *p-value<0.0001)*, financial ratios (SQ001) (Chi-square: 3.87 and *p-value=0.049)* were also significant in the experiment. Predictions based on another influencing factor, fundamental analysis (SQ002), were also correct with a quite high percentage (55.4%), but did not reach statistical significance (Chi-square: 2.063 and *p-value=0.151).*

For the three-month predictions the most accurate predictions included group feedback (SQ003) as the basis of the decision-making. Three-month predictions based on that were significantly been better than by chance (Chi-square: 27.473 and *p-value*<0.0001). For three-month prediction others influencing factors such as expert opinions (SQ007) (Chi-square: 6.434 and *p-value=0.011)*, financial ratios (SQ009) (Chi-square: 21.884 and *p-value<0.0001)* and other were also significant in the experiment.

The overall (average) most accurate predictions (57.0% correct) included expert opinions (SQ007), but technical analysis (54.8% correct) and fundamental analysis (54% correct) were also among the options of basic principles mentioned by the participants. Hence it can be concluded that the data from the experiments supports A14.

An additional observation regarding interpretation of the data is that experts themselves were not necessarily better. However, to use their expertise and include it in one's own reflective decision-making appears so aid predictive accuracy. This finding even holds for the experts themselves: the predictions where experts stated that they used information from (other) experts were considerably more accurate (61.1% correct) than their average recommendations (53.4% correct). Despite their comparably easy access to expert opinions they used the option to include (other) expert opinions (SQ007) only for 24.82%. The interpretation that reflection on other ideas and/or opinions increases predictive quality is also supported by the fact that participants who included group results in their own decision-making were also comparatively successful with their forecasts.

Area of Discussion A15: People who base their predictions on several decision-making approaches and/or information sources are better than those who make decisions based on fewer approaches/sources. Some decision-making approaches work better than others.

Further analysis of the participants' data provided to the question about the basis of their decision and influencing factors could also help to inform A15. The data on the use of information sources and basis for predictions were summed and a new sum variable (SQSUM) formed: the range of this variable is from 1 to 8 (number of information sources/decision-making approaches). The question is whether, with an increasing number of sources of information used, the number of correct predictions increases. As shown in the following table, the comparison shows considerable differences, in particular for the 1-week and 1-month periods (see table 91).

Table 91. *Overview of decision-making fundamentals and basic information*

	All Answers		Answers with rationale		SQSUM		Average (SQSUM / Answers)	
	correct	wrong	correct	wrong	correct	wrong	correct	wrong
1-week	2284	2289	2272	2272	6548	3640	2.8820422535	1.6021126761
1-month	2291	2282	2274	2270	5012	5176	2.2040457344	2.2801762115
3-month	2029	2544	2014	2530	5165	5023	2.5645481629	1.9853754941

The table (see table 91) shows that for the 1-week predictions respondents with correct predictions used on average about 2.88 sources of information and with false predictions about 1.60 less. This indicates that for short term predictions it might be an advantage to be informed by the media and/or to apply analytical techniques.

For the 1-month predictions almost no difference could be observed. Respondents with correct predictions used on average about 2.20 sources of information and with false predictions about 2.28 or even slightly more. Again for the 3-month predictions respondents with correct predictions used on average about 2.56 sources of information and with false predictions about 1.99 less. This indicates that for longer term predictions it might also be an advantage to be informed by the media and/or to apply analytical techniques. In summary the data from the experiments indicates that there is a benefit from using several decision-making approaches and/or information sources, i.e. assumption A15 is supported by the data.

But it is not only relevant to use several approaches: there are also differences in the accuracy of the predictions and the approach applied. Participants who provided answers (i.e., ticked a box in the questionnaire) about their approach had higher success rates with some approaches and lower success rates with other approaches as compared with participants not using this particular approach (see Tables 137, 138, and 139 in the appendix Annex IV). For the 1-week predictions significant differences in the frequencies of correct and wrong predictions (see Chi-Square test in Table 137) are especially prevalent with the following groups:

- Fundamental analysis [SQ002] (Participants had 4.4% correct predictions and 3.3% wrong; *p-value*=0.037)
- Group results [SQ003] (Participants had 7.1% correct predictions and 8.9% wrong; *p-value*=0.029)
- Market sentiment [SQ006] (Participants had 43% correct predictions, however 46.2% wrong; *p-value*=0.034)

For the 1-month predictions significant differences in the frequencies of correct and wrong predictions (see Chi-Square test in Table 138) are especially prevalent with the following groups:

- Financial ratios [SQ001]: Participants had 9.5 % correct predictions and 7.8 % wrong; *p-value=0.038*
- Experts [SQ007]: Participants had 12.4% correct predictions and 9.9% wrong; *p-value=0.007*
- Technical analysis [SQ009]: Participants had 16.7% correct predictions, however 12.3% wrong; *p-value<0.0001*

For the 3-month predictions significant differences in the frequencies of correct and wrong predictions (see Chi-Square test in Table 139) are especially prevalent with the following groups:

- Group results [SQ003]: Participants had 9.2% correct predictions and 6.6% wrong; *p-value=0.001*
- Intuition [SQ005]: Participants had 83% correct predictions and 86.6% wrong; p-*value=0.001*
- Market sentiment [SQ006]: Participants had 42.4% correct predictions, however 47.4% wrong; *p-value=0.001*
- Experts [SQ007] (Participants had 12.3% correct predictions and 9.6% wrong; *p-value*=0.005)

Area of Discussion A16: People are better at predicting steady upward or downward trends than changes of direction.

A comparison of the data of the experiment split into two groups showed that there is no significant impact on predictive accuracy whether the stock price has the same direction as the week before. In order to assess A16 the results of all participants were grouped into two categories (see Tables: 123, 128 and 129 in the appendix). First category: The direction of the stock price movement (up or down) is the same as in the week before, implicating a steady trend (see Table 92), and the second group with a different stock price development than the previous week (see Table 93). Missing values have been excluded in this analysis.

Table 92. *Prediction Quality with Intact Stock Price Trend*

	Trend intact				
		Correct	Wrong	Excluded	Correct (%)
1-Week Predictions	Sum	6509	6395	3911	50.12%
1-Month Predictions	Sum	8719	8711	5285	50.38%
3-Month Predictions	Sum	10125	7930	5545	55.31%

Table 93. *Predictive Quality with Stock Price Direction Different to previous Week*

	No trend				
		Correct	Wrong	Excluded	Correct (%)
1-Week Predictions	Sum	3726	3710	2299	49.56%
1-Month Predictions	Sum	1491	1419	925	51.18%
3-Month Predictions	Sum	1355	930	665	58.41%

Generally, data indicates that the assumption should be rejected. There was no statistical significance in the difference between both categories. A chi-square test of the 1-week predictions results in Chi-square 0.211 and *p-value: 0.65* and a chi-square test of the 1-month predictions results in Chi-square 1.47 and *p-value: 0.23*. The analysis of the 3-month predictions with a chi-square test showed a significant result (Chi-square 8.56 and *p-value*: 0.003). However, both the results with trend and even more without trend were significantly better than by chance, and the percentage rate of correct predictions (58.41%) in the "no trend" situation was better than prediction with an intact trend (55.31%). In summary all these results support the conclusion that A16 should be rejected.

Area of Discussion A17: Certain individuals are especially good at prediction, as compared with the average of other members of a given group, of which they are members.

The analysis of the data from the experiment showed that there were indeed a number of participants who appeared to be especially good at predicting stock prices, as compared with the average of other members of a given group, of which they are members (see also2.3 Decision-Making and Forecasting). Table 94 shows the ten most correct and ten least correct members of the main experiment. Obviously it is not necessary for participants to be particularly interested in the stock-market to provide correct predictions, nor do people seem to have a very good sense of their own abilities. There are quite a few people in the top ten with a Self-Assessment of their skill below 5, i.e. below average skills. On the other hand there are also quite a few people with a quite high Self-Assessment of their knowledge about stock-markets among the ten least correct participants. However, it is noticeable that participants with a PID-S plus score are over-represented among the best participants.

Table 94. *The Ten Most Correct and Ten Least Correct Participants*

Participant ID	Correct	Wrong	Excluded	Correct (%)	Group	PID Score	Emotional self-assess	Skill self-assess	Interested in stock market
204	360	115	25	75,8%	AG	PID-D	rather rational	10	yes
604	335	115	50	74,4%	EDG	PID-D	rather rational	4	yes
615	370	130		74,0%	EDG	PID-S plus	emotional	3.5	no
616	325	125	50	72,2%	EDG	PID-S minus	rational	4	no
503	355	145		71,0%	NFG	PID-S minus	rather emotional	6	no
510	265	110	125	70,7%	NFG	PID-S	rather	3	no

						plus	emotional		
621	300	125	75	70,6%	EDG	PID-S plus	rather emotional	2.5	no
511	335	140	25	70,5%	NFG	PID-S plus	rather emotional	4	no
102	225	100	175	69,2%	PG	PID-I	emotional	3	yes
Average				**54,8%**					
520	10	15	475	40,0%	NFG	PID-I	definitely emotional	5	no
501	190	310		38,0%	NFG	PID-S minus	rather emotional	6.5	yes
105	150	250	100	37,5%	PG	PID-D	rational	8	yes
515	75	125	300	37,5%	NFG	PID-D	emotional	3	no
617	185	315		37,0%	EDG	PID-S minus	rational	6	yes
203	160	315	25	33,7%	AG	PID-S plus	rational	9.5	yes
521	140	335	25	29,5%	NFG	PID-D	rather emotional	3	no
509	95	230	175	29,2%	NFG	PID-S minus	rational	7	yes
2	135	365		27,0%	IG	PID-D	rather rational	6	yes

All participants were interviewed to gain a better understanding of the differences between the best and the less good participants. A comparative analysis of the interview data of the best and least correct participants is provided in section “Comparative Analysis of Best and Worst Predictors” on page 305 and the discussion of assumption A21 on page 224.

Area of Discussion A18: Predictions by lay people of well-known shares, such as Adidas, are better than their predictions of lesser known (to the general public) shares.

In addition to individual characteristics or processing the information from various sources and decision-making approaches the identity of the companies themselves might influence correct or incorrect predictions, i.e., it may be the case that certain companies can be predicted more reliably than other companies.

It might be easier for participants to provide precise predictions if they are more familiar with the company and its products, hence predictions by lay people about well-known shares, such as Adidas, are better than their predictions about lesser known shares. Since predictions about different companies in different industries were gathered during the experiment assessment of accuracy for the different shares is possible. However, the design of the sampling for experiment purposely focused on well-known companies listed in the major German stock-index DAX. Accordingly, all companies are more or less well known to the participants. Thus the sampling implies a limitation to the assessment of A17. Nevertheless, it might be true that companies with retail focus, like Adidas, might be easier for lay people to predict. The data supports this assumption to some extent. Predictive quality was slightly better for Adidas compared with the average accuracy of the lay group. . The overall accuracy of the lay groups was best for the Adidas share price predictions. The lay groups combined (EDG, IG, and NFG) had an accuracy of 72.5% Correct compared with 47.7% on average for the other 4 shares combined.

Table 95. *Aggregated Main Run Predictions*

Adidas	AG	EDG	IG	NFG	PG	Expert	Measurements
Correct	18	45	32	46	32	35	60
Wrong	28	12	22	12	16	25	
Excluded	14	3	6	2	12	0	
Correct (%)	39.1%	78.9%	59.3%	79.3%	66.7%	58.3%	
Heidelberg Cement	AG	EDG	IG	NFG	PG	Expert	
Correct	15	27	22	31	33	49	60
Wrong	34	32	31	25	17	8	
Excluded	11	1	7	4	10	3	
Correct (%)	30.6%	45.8%	41.5%	55.4%	66.0%	86.0%	
RWE	AG	EDG	IG	NFG	PG	Expert	
Correct	48	19	32	32	24	33	60
Wrong	10	37	22	25	24	9	
Excluded	2	4	6	3	12	18	
Correct (%)	82.8%	33.9%	59.3%	56.1%	50.0%	78.6%	
Siemens	AG	EDG	IG	NFG	PG	Expert	
Correct	24	32	18	27	36	28	60
Wrong	23	25	35	27	15	29	
Excluded	13	3	7	6	9	3	
Correct (%)	51.1%	56.1%	34.0%	50.0%	70.6%	49.1%	
ThyssenKrupp	AG	EDG	IG	NFG	PG	Expert	
Correct	29	31	14	29	23	26	60
Wrong	27	28	39	31	29	28	
Excluded	4	1	7	0	8	6	
Correct (%)	51.8%	52.5%	26.4%	48.3%	44.2%	48.1%	

At all three time periods significant differences in the frequency distributions of correct and incorrect predictions could be observed, depending on the company. However, the differences are not the same direction, depending on the period. The share development in the 1-week periods of Siemens and ThyssenKrupp has significant differences in false and correct predictions, but not for the other companies. In the case of Siemens, the number of correct predictions is higher than the wrong predictions, at ThyssenKrupp the reverse. For Siemens the number of correct predictions is higher than the expected frequency (assuming a random distribution), though at ThyssenKrupp lower than expected. Each subscript letter in table 96, 97, and 98 indicates a subset of categories whose column proportions do not differ significantly from each other on the .05 level.

Table 96. *Crosstab Survey Share * 1 Week Predictions*

Crosstab Survey Share * 1-Week Predictions			1-Week Predictions		Sum
			wrong	correct	
SurveyShare	Adidas	Frequency	442_a	473_a	915
		Expected Frequency	458.0	457.0	915.0
		% in 1-Week Predictions	19.3%	20.7%	20.0%
	HeidelbergCement	Frequency	484_a	431_a	915
		Expected Frequency	458.0	457.0	915.0
		% in 1-Week Predictions	21.1%	18.9%	20.0%
	RWE	Frequency	466_a	449_a	915
		Expected Frequency	458.0	457.0	915.0
		% in 1-Week Predictions	20.4%	19.7%	20.0%
	Siemens	Frequency	406_a	508_b	914
		Expected Frequency	457.5	456.5	914.0
		% in 1-Week Predictions	17.7%	22.2%	20.0%
	ThyssenKrupp	Frequency	491_a	423_b	914
		Expected Frequency	457.5	456.5	914.0
		% in 1-Week Predictions	21.5%	18.5%	20.0%
Sum		Frequency	2289	2284	4573
		Expected Frequency	2289.0	2284.0	4573.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	20.873[a]	4	.000
Frequency of valid Data Sets	4573		

Predictive accuracy for the share development of Adidas and Heidelberg Cement in the 1-month periods had also significant differences in false and accurate predictions, but not for the other companies. For Adidas the number of correct predictions is higher than the wrong ones, for HeidelbergCement there are more wrong predictions. This implies that for Adidas the number of correct predictions is significantly higher than the expected number (assuming a random distribution), for HeidelbergCement, however, significantly lower.

Table 97. *Crosstab Survey Share * 1 Month Predictions*

Crosstab Survey Share * 1-Month Predictions			1-Month Predictions		Sum
			wrong	correct	
SurveyShare	Adidas	Frequency	392_{a}	523_{b}	915
		Expected Frequency	458.4	456.6	915.0
		% in 1-Month Predictions	17.1%	22.9%	20.0%
	HeidelbergCement	Frequency	567_{a}	348_{b}	915
		Expected Frequency	458.4	456.6	915.0
		% in 1-Month Predictions	24.7%	15.2%	20.0%
	RWE	Frequency	444_{a}	471_{a}	915
		Expected Frequency	458.4	456.6	915.0
		% in 1-Month Predictions	19.4%	20.6%	20.0%
	Siemens	Frequency	451_{a}	463_{a}	914
		Expected Frequency	457.9	456.1	914.0
		% in 1-Month Predictions	19.7%	20.3%	20.0%
	ThyssenKrupp	Frequency	437_{a}	477_{a}	914
		Expected Frequency	457.9	456.1	914.0
		% in 1-Month Predictions	19.1%	20.9%	20.0%
Sum		Frequency	2291	2282	4573
		Expected Frequency	2291.0	2282.0	4573.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	73.859[a]	4	.000
Frequency of valid Data Sets	4573		

Predictive accuracy for the share development of Adidas, RWE, Siemens and ThyssenKrupp has significant differences in the number of wrong predictions and correct, but not for HeidelbergCement. For Adidas, the number of correct predictions is higher than the wrong, RWE is also higher, for Siemens and ThyssenKrupp in contrast lower. For Adidas the number of correct predictions is significantly higher than the expected number (assuming a random distribution), but for RWE lower. The same is true for Siemens and ThyssenKrupp: the number of correct predictions is lower than the expected number.

Table 98. *Crosstab Survey Share * 3 Month Predictions*

Crosstab Survey Share * 3-Month Predictions			3 Month Predictions		Sum
			wrong	correct	
Survey Share	Adidas	Frequency	265_a	650_b	915
		Expected Frequency	406.0	509.0	915.0
		% in 3-Month Predictions	13.1%	25.6%	20.0%
	HeidelbergCement	Frequency	395_a	520_a	915
		Expected Frequency	406.0	509.0	915.0
		% in 3-Month Predictions	19.5%	20.4%	20.0%
	RWE	Frequency	447_a	468_b	915
		Expected Frequency	406.0	509.0	915.0
		% in 3-Month Predictions	22.0%	18.4%	20.0%
	Siemens	Frequency	464_a	450_b	914
		Expected Frequency	405.5	508.5	914.0
		% in 3-Month Predictions	22.9%	17.7%	20.0%
	ThyssenKrupp	Frequency	458_a	456_b	914
		Expected Frequency	405.5	508.5	914.0
		% in 3-Month Predictions	22.6%	17.9%	20.0%
Sum		Frequency	2029	2544	4573
		Expected Frequency	2029.0	2544.0	4573.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	123.338[a]	4	.000
Frequency of valid Data Sets	4573		

Area of Discussion A19: Financial analysts perform better in the private setting compared with their public forecasts.

There is a long and ongoing discussion in the literature about the influence of the design of incentive schemes for forecasters (Osband, 1989) and, more specifically, influence factors like consensus-seeking, competition and publicly stated and private (non-public) forecasts (Batchelor & Dua, 1992; Lichtendahl, Grushka-Cockayne, & Pfeifer, 2013; Marinovic & Ottaviani, 2013; Ottaviani & Sørensen, 2006). This is particularly relevant for financial analysts when incentives (Aiolfi et al., 2009; Beyer & Guttman, 2011) and conflicts of interest may influence their publicly stated forecasts (Bolliger, 2004, 2004; Lin & McNichols, 1998; Stanzel, 2007). Hence, it might be hypothesised that financial analysts would perform better in the private setting compared with their public forecasts (Endress, 2014).

This section presents findings from the data generated using financial analysts' stock price predictions in the private setting of the experiment and their publicly stated forecasts as published on Bloomberg. The aim is to assess the effect of publication pressure and group dynamics on stock price predictions, to assess whether there is a difference between non-public and published recommendations and to identify the underlying key mechanisms of the decision-making process.

The financial analysts provided 60 correct answers out of 90 measurements (data for 10 measurements from the total of 100 measurements were missing) in private—that is, the estimates given anonymously during the main run of the experiment. Compared with the open data published on Bloomberg, this is a considerably lower number of correct answers. Seventy-six of the answers were from the open 100 measurements on Bloomberg. The target price accuracy was considerably higher in the private setting. The overall higher target price accuracy was higher in the experiment compared with their open price targets published on Bloomberg. However, not all the analysts' target price estimations were more accurate.

Table 99. *Analysts' Private and Open Target Price Accuracy from the Main Experiment*

	Experiment (Private) Target Price Accuracy	Bloomberg (Open/Public) Target Price Accuracy
Adidas	6.60%	8.03%
HeidelbergCement	5.17%	6.34%
RWE	15.31%	4.58%
Siemens	7.95%	27.70%
ThyssenKrupp	18.41%	11.21%
Average	11.07%	11.57%

During the main experiment, it could be observed that financial analysts were slightly less optimistic in the open setting compared with the private experiment. The difference between the public and private recommendation is not significant (Chi-square: 0.098; *p-value*=0.754). The analysts provided 60% (60) buy and 40% (40) sell/avoid recommendations in the open setting, compared with only 62.2% (56) and 37.8% (34) sell recommendations in private (10 missing private measurements have been excluded). However, this difference is very small and possibly affected by the news flow and situation with Siemens (e.g. Höhler, 2013; Rubenfeld, 2013) and ThyssenKrupp (e.g. Ott, 2012; Sheahan, 2013) involved in issues with very negative sentiment. Regulatory issues, particularly the changes in German energy policy, also provided a quite negative sentiment for RWE (Eckl-Dorna, 2013) which might have influenced the public perception of the company. It is still possible that this market environment had an influence on the results. Further research is needed to test the assumption of overoptimistic public recommendations by financial analysts.

The main experiment indicated that experts were more likely to change recommendations in the private setting. While the analysts did not change their open recommendations on Bloomberg during the 10-week period of the main experiment, there were six changes of recommendations in the private setting. Additionally, there was only one change in target price on Bloomberg, but there were 70 changes in price target in the private setting. In addition, 26 missing private measurements were excluded, and the target price was unchanged only four times. Obviously, the analysts were much more likely to change their opinions in the anonymous setting. That might have contributed to the higher price target accuracy.

Some preliminary findings from the interviews are that almost all professionals mentioned that they are very interested in the markets and even see it as a hobby. Most of them trade shares themselves on a private account. One analyst stated: “I trade stocks myself. That is probably the strongest argument. It's a hobby of mine. I read a lot that has to do with or could have to do with it” or “It manifests itself first of all in that I also invest my private money in stocks, that is also in single instruments, and also that I inform myself just before I do that. This is on the one hand, of course, a purely financial investment, but on the other hand also interesting. In a sense it is also a hobby.” When answering the question as to how they made their decision, the professionals still referred to intuition and gut feelings. They included answers such as the following: “Frequently market climate and intuition and gut feeling, probably even more than valuation, although I know the valuation and ratings of companies that I do not cover as well.” But it seems that gut feelings, as well as other, similar initial classifications, are not the same for professionals, who described a quite different decision approach. When asked to describe their intuition and gut feelings, they answered as follows: “[It] has a lot to do with the development of the stock in the last few days and how I generally estimate the market. So for example, I guess the market is not so great, and the stock previously went very well, then I guess it's not so good, it's probably going down” or “Yes, more like the general market sentiment, the news flow, macro but also micro, so to speak, and how I perceive it, so that's not carefully analysed but rather the current mood.”

The analysis indicated that even for professional financial analysts who usually describe themselves as rather rational people, decisions about investments or stock price recommendations are not always very consistent and, amongst other factors, are influenced by emotional factors. Slightly different settings and ways of framing the questions have considerable impact on the decisions made. The data collected supports, to some extent, the idea that the incentives for analysts and public competitions might induce financial analysts to report strategically (Lichtendahl et al., 2013) and that might lead, in some cases, to a reduced quality of recommendations. There are indications that anonymously given forecasts might be better, because analysts do not incur any peer pressure or incentives and there is no need to justify any changes in their opinions, but this was only found with target price predictions. Still, there were not always better results with the anonymous/private setting. In the main run, it was found that there was not only a higher rate of recommendation change activity but that the public recommendations were actually considerably better than the private ones.

A completely different result was found with the target prices. There was a much higher change activity amongst these prices. The analysts provided, in almost every round, new target prices. In this case, the private estimates of target prices led to considerably higher accuracy of target price forecasts for most stocks. That might be partly explained by the significantly higher number of changes in price recommendations. Further research might help to gain a more holistic understanding of the decision-making process and to create an explanatory schema.

Nevertheless, it might be useful to conduct further experiments with other market conditions, different stocks, and variations of the questionnaire design. With more data, it might be possible to gain a better understanding of questions like the following: Are financial analysts consistently overoptimistic in their forecasts of covered stocks? Or are analysts better than laypeople in bull markets but lose that advantage in bear markets? This could not be fully addressed with the experiments conducted. Additionally, the data from the in-depth interviews might provide some more information about the factors influencing the underlying decision-making process.

Area of Discussion A20: The same lay people who are good at short term predictions are also good at longer term predictions, as compared with the group average.

There are a few participants who were within the 20 participants with most correct predictions and in the top 20 for 1-month predictions and top 20 for 3-month predictions. However, the number of participants does not significantly exceed the number of participants expected to be within the top list just by chance. Only two participants were in the top 20 for predictions for all three time-frames. While the best 20 participants for one-week predictions (averaging 57% correct answers) are on average still slightly better than the overall average for one-month predictions and three-month predictions (see Table 100), the reverse control for the best 20 participants for three-month predictions (averaging 68.1% correct answers) are even slightly less accurate than the overall average (see Table 101). Participants who performed well for periods other than the primary sort criteria of the table are highlighted in green.

Table 100. *Overview Best 20 Participants for a One Week Period*

PID	1W correct (%)	1-Week Rank	1M correct (%)	1-Month Rank	3M correct (%)	3-Month Rank	All correct (%)	Overall Rank
501	67.0%	1	49.0%	29	51.4%	37	45.2%	49
204	62.1%	2	71.6%	1	38.0%	52	51.3%	28
516	60.0%	3	65.3%	6	48.6%	40	52.4%	25
605	60.0%	4	45.7%	39	63.3%	18	59.4%	10
509	58.5%	5	41.5%	48	58.8%	28	44.2%	51
502	57.1%	6	51.4%	21	41.7%	50	46.7%	44
511	56.8%	7	62.1%	9	74.4%	2	60.4%	5
203	56.8%	8	45.3%	40	42.2%	49	48.9%	37
604	56.7%	9	50.0%	25	27.0%	59	37.7%	57
508	56.4%	10	65.5%	5	29.2%	58	43.1%	53
613	56.3%	11	58.8%	14	49.0%	39	47.3%	41
503	56.0%	12	68.0%	3	61.1%	21	62.1%	4
514	56.0%	13	49.0%	30	65.0%	14	60.0%	7
513	56.0%	14	37.0%	54	58.9%	26	46.7%	43
512	56.0%	15	53.3%	20	45.6%	44	49.6%	34
201	54.3%	16	61.4%	11	58.8%	27	57.9%	14
519	54.0%	17	55.0%	19	61.1%	20	57.0%	15

105	53.8%	18	41.3%	49	51.4%	38	48.6%	38
618	53.3%	19	56.7%	17	60.0%	24	56.3%	18
205	53.3%	20	50.0%	26	70.7%	6	59.6%	9
1-Week Best 20 Avg.	57.0%	10.5	53.9%	23.3	52.8%	32.6	51.7%	29.1
Overall Avg.	49.3%		50.1%		54.8%		51.4%	

Generally the data indicates that assumption A20 should be rejected. An assessment of the correlation[2] between the predictive accuracy of the one-week and the one-month predictions indicated a weak positive relationship (0.237). The correlation between the predictive accuracy of the one-week and the three-month predictions indicated a moderate negative relationship (-0.300). The correlation between the predictive accuracy of the one-month and the three-month predictions indicated no relationship or a negligible one (-0.012).

Table 101. *Overview Best 20 Participants for a Three Month Period*

PID	1W correct (%)	1-Week Rank	1M correct (%)	1-Month Rank	3M correct (%)	3-Month Rank	All correct (%)	Overall Rank
620	46.7%	39	68.3%	2	75.8%	1	0.6982	1
511	56.8%	7	62.1%	9	74.4%	2	60.4%	5
515	30.0%	59	45.0%	43	74.0%	3	56.7%	16
610	38.9%	54	42.2%	46	72.2%	4	49.6%	33
36	38.2%	55	60.0%	12	71.0%	5	65.0%	2
205	53.3%	20	50.0%	26	70.7%	6	59.6%	9
608	41.0%	52	41.0%	51	70.6%	7	53.3%	22
603	50.0%	30	65.0%	7	70.5%	8	63.2%	3
507	37.1%	56	42.9%	45	69.2%	9	58.5%	11
621	42.4%	51	47.1%	36	68.9%	10	60.4%	6
520	40.0%	53	40.0%	52	68.0%	11	50.0%	32
506	31.3%	58	51.3%	23	67.8%	12	55.9%	19
518	52.5%	22	50.0%	27	66.0%	13	51.7%	26
514	56.0%	13	49.0%	30	65.0%	14	60.0%	7
615	45.0%	43	51.0%	24	64.0%	15	58.3%	13
611	49.0%	32	41.0%	50	63.8%	17	51.7%	27
4	42.9%	49	51.4%	22	63.8%	16	55.4%	21
605	60.0%	4	45.7%	39	63.3%	18	59.4%	10
2	51.0%	25	35.0%	55	61.8%	19	53.3%	23
519	54.0%	17	55.0%	19	61.1%	20	57.0%	15

2 Correlation assessment according to: http://faculty.quinnipiac.edu/libarts/polsci/Statistics.html

3 Month Best 20 Avg.	45.8%	37.0	49.6%	30.9	68.1%	10.5	57.5%	15.1
Overall Avg.	49.3%		50.1%		54.8%		51.4%	

While different participants performed well for the different periods, it is still possible to find similarities between the top performers. It is conspicuous that for all periods people with the PID-Score "PID-S plus" were over-represented (see Table 102). Furthermore, 7 of the 9 participants (77.8%) with "PID-S plus" were within the 25 best predictors for a three-month period.

Table 102. *Relative Proportion of the PID-Scores among the Top Twenty Predictors*

1-Week Predictions	Proportion within the Top 20 Participants	Total Number	Relative Proportion
PID-I	4	11	36.4%
PID-D	7	22	31.8%
PID-S minus	5	17	29.4%
PID-S plus	4	9	44.4%
1-Month Predictions			
PID-I	4	11	36.4%
PID-D	8	22	36.4%
PID-S minus	3	17	17.6%
PID-S plus	5	9	55.6%
3-Month Predictions			
PID-I	4	11	36.4%
PID-D	6	22	27.3%
PID-S minus	6	17	35.3%
PID-S plus	4	9	44.4%

Area of Discussion A21: If a good predictor is defined by having a higher value of number of correct predictions divided by number of incorrect predictions, then what are the characteristics of these good predictors, as found from the questionnaire results?

A total of 31 participants (out of 59 participants) had a predictive quality of more than 50% correct predictions. At first glance, it appears that the top predictors don't have a lot in common (see Table 103). However, an in depth-analysis reveals some striking patterns. As well as the facts already discussed, e.g. that top predictors have an above average educational level (74.2% of the top predictors have an academic degree compared with 67.8% of the overall participants, see also A5, p. 166) and are over-represented in the PID-S-plus category (see also A17, p. 208) it was interesting and conspicuous that quite a few of the top predictors mentioned during the interviews that they used news (online, newspaper, company news etc.), but that their interpretation of the news was intuitive. Typical responses by the good predictors to the question in the interview about the bases for their decisions were “that was definitely the latest news or economic decisions. I've always kept in mind that it's winter and they won't sell as much, because the economy does not pick up and the construction business is the engine of the economy, but that will catch up later. Sportswear, Adidas always sells, Christmas business is always good, economy is great. ThyssenKrupp is a “no show” at the moment, so is somehow difficult. What about RWE, the energy transition? I strongly based it in the political events of the day, which was actually my anchor.” or “It depends. Now at the moment rather emotional things. Some are based on news, where you connect one thing with another, but I think more the emotional side of stories, where I thought that affects such and such. […]”.

Table 103. *Participants with 50% or more correct predictions.*

ID	Interested in Stock Market	Skill	Univ. Degree	Age	PID-D	PID-I	PID_Score	Emotional Self-Assessment	Pred. Accu. (ALL)	Rank (all)
204	yes	10	1	31	4.22	2.56	PID-D	Rather rational	69.8%	1
503	no	6	1	46	3.67	2.89	PID-S minus	rather emotional	65.0%	2
511	yes	4	1	29	4.33	3.33	PID-S plus	rather emotional	63.2%	3
516	yes	7.5	1	43	4.33	2.78	PID-D	Rather rational	62.1%	4
604	yes	4	1	44	4.78	2.56	PID-D	Rather rational	60.4%	5
101	yes	9	1	41	3.56	2.33	PID-S minus	Rational	60.4%	5
603	No answer	No answer	1	58	4.22	3.00	PID-D	No answer	60.0%	7
508	yes	3	1	33	4.00	2.67	PID-D	Rather rational	60.0%	7
510	yes	3	1	33	4.00	3.33	PID-S plus	Rather emotional.	59.6%	9
620	no	1		43	3.44	2.89	PID-S minus	rational	59.4%	10
601	yes	5.5	1	38	3.67	3.00	PID-S minus	Rather rational	58.5%	11
102	yes	3	1	31	3.67	3.67	PID-I	Emotional	58.5%	11
517	no	2.5	1	38	3.44	2.67	PID-S minus	In-between	58.3%	13
613	yes	4.5	1	31	4.67	3.33	PID-S plus	Emotional	57.9%	14
618	yes	4.5	1	41	3.56	3.33	PID-I	Rather rational	57.0%	15
615	no	3.5		29	4.33	3.33	PID-S plus	Emotional	56.7%	16
512	no	1		37	3.00	3.56	PID-I	Emotional	56.4%	17
519	no	2.5		41	3.67	4.00	PID-I	Emotional	56.3%	18
614	no	3		26	4.00	3.11	PID-D	Rather rational	55.9%	19
201	yes	8	1	33	4.00	2.89	PID-D	rational	55.7%	20
518	yes	2.5	1	51	4.22	2.22	PID-D	Rational	55.4%	21
36	yes	6.5	1	42	3.89	3.67	PID-S plus	In-between	53.3%	22
621	yes	3	1	35	4.44	3.44	PID-S plus	Rather emotional	53.3%	22
514	yes	1		58	2.78	2.78	PID-S minus	Very emotional	52.7%	24
502	No answer	4	1	46	4.33	3.56	PID-S plus	Emotional	52.4%	25
606	yes	6	1	38	3.22	2.89	PID-S minus	Rather emotional	51.7%	26
607	yes	2		58	3.11	3.78	PID-I	Emotional	51.7%	26
501	yes	6.5	1	41	3.11	3.11	PID-S minus	rather emotional	51.3%	28
38	yes	5.5	1	43	3.89	2.44	PID-D	Rather rational	50.7%	29
202	yes	9	1	33	4.00	2.67	PID-D	Emotional	50.7%	29
605	No answer	4		38	1.56	3.56	PID-I	Rather rational	50.5%	31

5.8.2 Findings of the main experiment

The findings of the main experiment did not generally confirm that groups of lay people are better at predicting stock price movements than the experts. Still, it might be noteworthy that the best performance for the one-week predictions in the main experiment was by the non-feedback group (NFG), while the best performance for the one-month predictions in the main experiment was by the expert analysts (individual experts) and for the three-month predictions from the financial professionals (PG). Generally, it can be seen that lay groups (EDG, IG, and NFG) did not perform per se better than the professionals (AG, PG, and individual experts), but that the groups without a feedback loop performed slightly better: this was particularly true for the one-week and one-month predictions. This finding agrees with the suggestion that diverse and independent decision-making by individual group members is supportive of collective intelligence (Page, 2008b).

The analysis of the in-depth qualitative interviews delivered insight into the decision-making process of the participants and revealed that for many participants financial predictions are a rather emotional issue. Most participants, lay people and experts, stated that intuition played an important role for their stock price predictions within the experiment. However, there are differences in the “use of” intuition or “gut feeling”. While it appears that their intuition is for poor predictors like a “random guess” leading to rather thoughtless statements, it seems to be the case that the good predictors base their intuition on several factors—including fundamental and macroeconomic considerations.

To complete the analysis of data from the main experiment a total of 21 assumptions have been tested. Table 104 shows a summary of the test results. The rich data set of the experiment allowed a quite extensive assessment of many aspects. Clearly, not all the assumptions formulated are supported by the data from the experiment. About half of the assumptions are not supported.

Table 104. *Discussion Summary*

Areas of discussion	**Supported by main experiment**
A1 A lay person may better predicting short term (1 week) than a professional financial analyst, but over a longer period the analysis models of an analyst will lead to better results.	not supported
A2 There is an improvement in predictive accuracy results from feedback from an e-Delphi group.	not supported
A3 Predictive quality improves over time as people learn about the stocks.	supported
A4 Rational people are better at financial decision-making compared with intuitive people	not supported
A5 Educational level has a major impact on the ability to predict stock prices	supported
A6 Female persons are better at assessing market sentiment and outperform with short term predictions.	supported
A7 Financial analysts are consistently overoptimistic in their forecasts of covered stocks.	not supported
A8 Life experience and age have an influence on stock price predictions. Older people are more risk averse.	not supported
A9 Analysts are better than lay people in bull markets, but lose that advantage in bear markets.	not supported
A10 People who are interested in the stock market are able to provide better predictions.	not supported
A11 People who think that they know more about the stock market are able to make better forecasts.	not supported
A12 Predictions with a higher level of confidence are generally better than predictions with low confidence.	supported
A13 When people express a higher upside or downside (in terms of price to target price difference) the predictions are better.	supported
A14 Predictions based on fundamental (or technical?) analysis are superior to intuitive predictions.	supported
A15 People who base their predictions on several decision-making approaches and/or information sources are better than those who make decisions based on fewer approaches/sources.	supported
A16 People are better at predicting steady upward or downward trends than changes of direction.	not supported
A17 Certain individuals are especially good at predicting, as compared to the average of other members of a given group, of which they are members.	supported
A18 Predictions by lay people of well-known shares, such as Adidas, are better than their predictions of lesser known shares (to the general public).	supported
A19 Financial analysts perform better in the private setting compared with their public forecasts.	supported
A20 The same lay people who are good at short term predictions are also good at longer term predictions, as compared to the group average	not supported
A21 If a good predictor is defined by having a higher value of number of correct predictions divided by number of incorrect predictions, then what are the characteristics of these good predictors, as found from the questionnaire results?	See page 224

6 Conclusion and Contribution to Knowledge and Business Practice

The contribution to knowledge and business practice of this thesis may be classified as contributing in three categories. Primarily it informs the understanding of group decision-making in online groups, in particular groups which focus on financial predictions. Secondly, a research framework was developed to assess the quality of decisions in online groups. And thirdly it provides insight into personal decision-making and contributes to an understanding of factors that may improve the decision-making quality of financial decisions and predictions. A limitation of this analysis is the focus on Germany, the limited set of companies and participants, and duration of the experiments.

The results from the experiments indicate that, there is always a degree of 'random walk' that influences the prediction of stock prices. This is true for professional financial analysts, but also for online groups with a collective intelligence approach. The predictability of stock prices remains rather limited. However, there are factors that might help to improve predictive accuracy and to facilitate these factors may contribute to a superior decision-making process. Some factors are inherent in the personality of the predictor. The findings from the study indicate that intuition plays a significant role in the decision-making process not only for lay people, but also for financial analysts and other financial professionals. Also, there are observable differences in the intuitive decision-making of lay people and experts. Further data might help to bring additional clarification to these differences of the underlying process.

6.1 Group Decision Making in Online Groups

The experiment provided a rich data set with qualitative and quantitative components. The analysis of these data sets provided unprecedented insights into the decision-making process and predictive qualities of equity forecasts from online groups (to simulate collective intelligence) and professional equity analysts. The research study indicated that equity predictions by Internet groups are not per se superior to predictions by professional equity analysts. The absolute performance of the predictions of online groups [research question 1] in the data set was slightly above the expected value (assuming random distribution), but not statistically significant. The relative performance—compared with the recommendations from professional equity analysts—was inconsistent. The predictions by the equity analysts in the main experiment were significantly better than the aggregated lay group predictions. The experiment demonstrated that collective intelligence is not a panacea for equity predictions. However, discussions in online groups might contribute to the interchange of ideas and formation of opinion of (potential) investors. Several participants confirmed in the interview that they consider online communities to be useful (e.g. to counterbalance information from the bank). Still, most lay participants stated that they are not interested and would not participate. It appears that there is a certain self-selection of participants in these Internet groups and a gathering of peers. That might be one reason for the weak performance of these online groups, because group diversity is reported to be an important factor for collective intelligence (Page, 2008b).

In the Online experiment group learning, in terms of the effect of the feedback loop from the e-Delphi process on predictive quality, [research question 2], was not very strong. No systematic effect in terms of group learning among members of the groups with feedback loop could be observed in the survey. Additionally, the analysis of secondary data gathered in existing communities which focus on equity investments and experience from initial “real money” investment products also underlined the indication that Internet groups are not per se superior to financial professionals. Furthermore, the analysis of the existing communities revealed a number of practical problems and shortfalls of collective intelligence approaches in existing business designs (see section 2.6 Analysis of Existing Online Communities and Published Analyst Recommendations). The insight from this experiment may also contribute, inspire and help to enhance the business practice of existing Internet groups as well as collective intelligence approaches in the financial services sector in general, in particular to avoid false promises and unwarranted expectations. Predictions based on a group decision following a simple approach like the e-Delphi based online experiment and the existing online communities with focus on equity predictions examined in the study need to be handled with care. However, there are some indications that groups may provide added value in the context of equity predictions. The approach with smaller, smarter crowds (Goldstein et al., 2014; Mannes et al., 2014) seems to be a particularly promising approach. Identifying the ‘right members’ for the groups is a key factor. Further research would help to give a fuller picture. Nevertheless, the findings from this research project may help to identify superior predictors, even without a track record of predictions, but based on personality traits.

6.2 Research Framework to Assess the Quality of Decisions in Online Groups

There is also a methodological contribution to knowledge from this research project. A research framework has been developed to assess the quality of decisions in online groups. Based on a mixed method approach it allows one to identify the quality and underlying mechanisms of online-groups. It was focused on predictions (in particular equity predictions), but it may be adapted to a wide range of decisions. The research framework developed is a comprehensive framework including a set of tools and techniques, especially questionnaires (for quantitative and qualitative assessment), coding conventions, and analytic methods. The research framework allows the repetition of the study in other different cultural contexts to allow the identification of cultural differences, other investment instruments or with different groups of participants to assess the impact of certain personality characteristics on decision-making.

A few of the artefacts developed are newly created (e.g. the online survey, the questionnaire for the semi-structured interviews), while some others are adapted from other research studies (e.g. the PID-scale) or have refined existing instruments (e.g. the coding conventions). The pilot run as well as the main experiment have demonstrated that this research framework works, and provides an effective structure for research projects concerning online group decisions.

6.3 Insights for Individual Decision Making

These results are both encouraging and exciting: as well as the insight into group decision-making, there is also a contribution to knowledge in terms of new findings in the context of the individual decision-making of the participants [research question 3]. There are different variables, in terms of the individual characteristics of the participants, which indicated significant impact on the quality of equity predictions. These are in particular educational level, gender, PID-score, but there are also relevant variables that are related to a particular prediction like confidence, upside or downside potential, and decision-making approach (see Figure 26). While there is no single factor that makes a good predictor, it can be concluded there are many factors that influence the quality of equity predictions. Flexibility and scrutinizing the sources and approaches appears to be a key for good investment decisions.

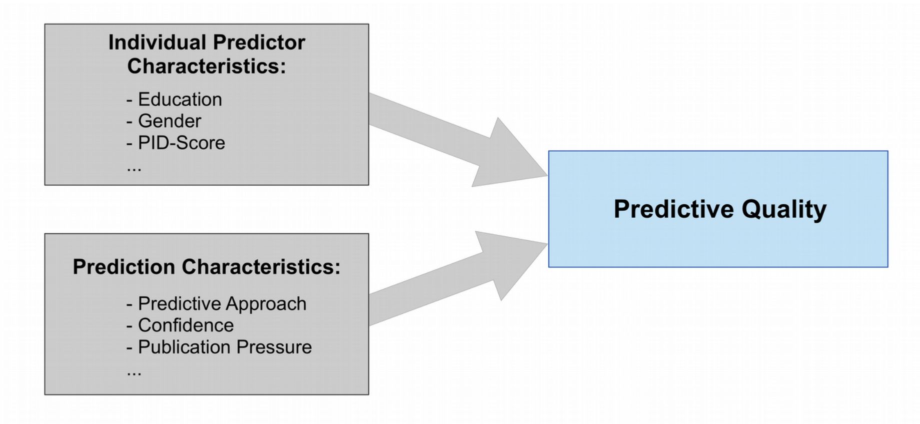


Figure 26: Influencing Factors for Share Price Predictive Quality

Participants—including top predictors—mentioned the ideal of rational decisions and that emotions are a not good context for financial decisions. Following this idea that rational people have a particular advantage, rational people would be expected to end up with the highest number of correct answers. Kahneman and Klein document this idea with a striking example:

> *[I]t is very likely that there are early indications that a building is about to collapse in a fire or that an infant will soon show obvious symptoms of infection. On the other hand, it is unlikely that there is publicly available information that could be used to predict how well a particular stock will do—if such valid information existed, the price of the stock would already reflect it. Thus, we have more reason to trust the intuition of an experienced fireground commander about the stability of a building, or the intuitions of a nurse about an infant, than to trust the intuitions of a trader about a stock. We can confidently expect that a detailed study of how professionals think is more likely to reveal useful predictive cues in the former cases than in the latter* (2009, p. 520).

Contrary to this example, the data gathered this experiment suggests that intuition is an important variable in the context of share price predictions. A more differentiated picture is probably necessary. In summary, it can be concluded that the assumption that public information can not be helpful to predict how well a stock will do is not supported by the data gathered in the experiment. However, the data gathered in the experiment indicated that intuitive people may have a slight advantage in terms of predictive quality. This means that an inverted version of the assumption would be correct: Intuitive people are better at financial decision-making compared with rational people. This is supported by the data, although the data are not significant in a direct comparison of PID-I and PID-D results at a level of significance of 0.05. Just to base decisions on intuition might not be the best strategy either. Still, in a domain where no one has complete information—like stock markets—purely rational approaches are not per se superior either. A combination of different approaches appears to be a superior strategy. This assumption is also supported by the data from the experiment. In particular, this can be documented by the PID-score analysis. A higher significance could be observed for the direct comparison of all four categories. The predictions of PID-S-plus participants are apparently of significantly higher accuracy. Still, these findings are from an experiment with a limited number of participants and should be repeated with a larger sample size and in different settings. Further research might be helpful to gain a better understanding of possible limitations and the underlying mechanisms.

There might be also a contribution to business practice in terms of the self-awareness, assignment and hiring of financial analysts. At the moment it appears that recommendations of financial analyst are based on rational models only. This is also reflected in the personality of the financial analysts. Most financial professionals in the panel (assessed by PID-score, interview and self-assessment of the participants) are rather rational and/or deliberative people. 60% of the financial professionals are type PID-D and 70% consider themselves to be rather rational or rational. However, the findings of this experiment suggests that people with a preference for <u>both</u> intuition and deliberation (type PID-S-plus) have an edge in terms of correct equity predictions. It might be worth considering the impact of this indication for further enhancements of the analysis as well as the hiring process for equity analysts. The findings of the experiment not only underline the fact that “intuition is an important component of professional competence in the domain of stock market” (Harteis & Gruber, 2008, p. 83), but also accentuates the need for a combination of deliberative and intuitive approaches. Additionally, the study has identified variables and provided indications that may help to enhance or establish guidelines and tools for business practice. The PID-Scale framework and findings of the experiment could contribute to building a sound foundation for the improvement of investment processes. Further research might be helpful to confirm the findings and allow inference for a wider range of settings and conditions.

Explanatory Model: Deliberated Intuition Model

The findings of the research project suggest a theory of the prediction process based on ‘deliberated intuition’, defined here as a considered decision to adopt an intuitive approach to making a prediction of share movement after reaching a limit for rational analysis. The factors influencing this considered decision are shown in figure 1, comprising the personality traits of the individual, their individual experience and training, and the situation in terms of risk and social context. This model emerged from reflection on the compiled analysis results, interview data and triangulation. The proposed ‘deliberated intuition model’ combines intuitive and deliberative elements and suggests three clusters of antecedents of the quality of prediction. This model of process suggests that prediction is based on conscious processing of intuition as a deliberate intention (Price & Norman, 2008). The suggestion that considerable time and effort are always expended when making predictions is clearly implausible, even if feasible for reaching a ‘better’ outcome. The concept of ‘bounded rationality’ (Simon, 1955) is compatible with ‘deliberated intuition’, implying that the decision to decide ‘*enough is enough*’ will depend on personal traits and perceived risks, coupled with social context, training and experience. The model of deliberated intuition which is proposed here views prediction as a process of practice which will be different for each individual. It is not a dual-processing model with two modes of process: rational deliberation or intuitive processes (eg. Tversky & Kahneman, 1974, 2002). It is a model which proposes that a predictor will decide, consciously or semi-consciously, when they feel ready to rely on gut-feeling, or to undertake more analysis. '*Interestingly, the degree to which people process information deliberately or intuitively has been found to depend on affective states*.' (de Vries, Holland, & Witteman, 2008).

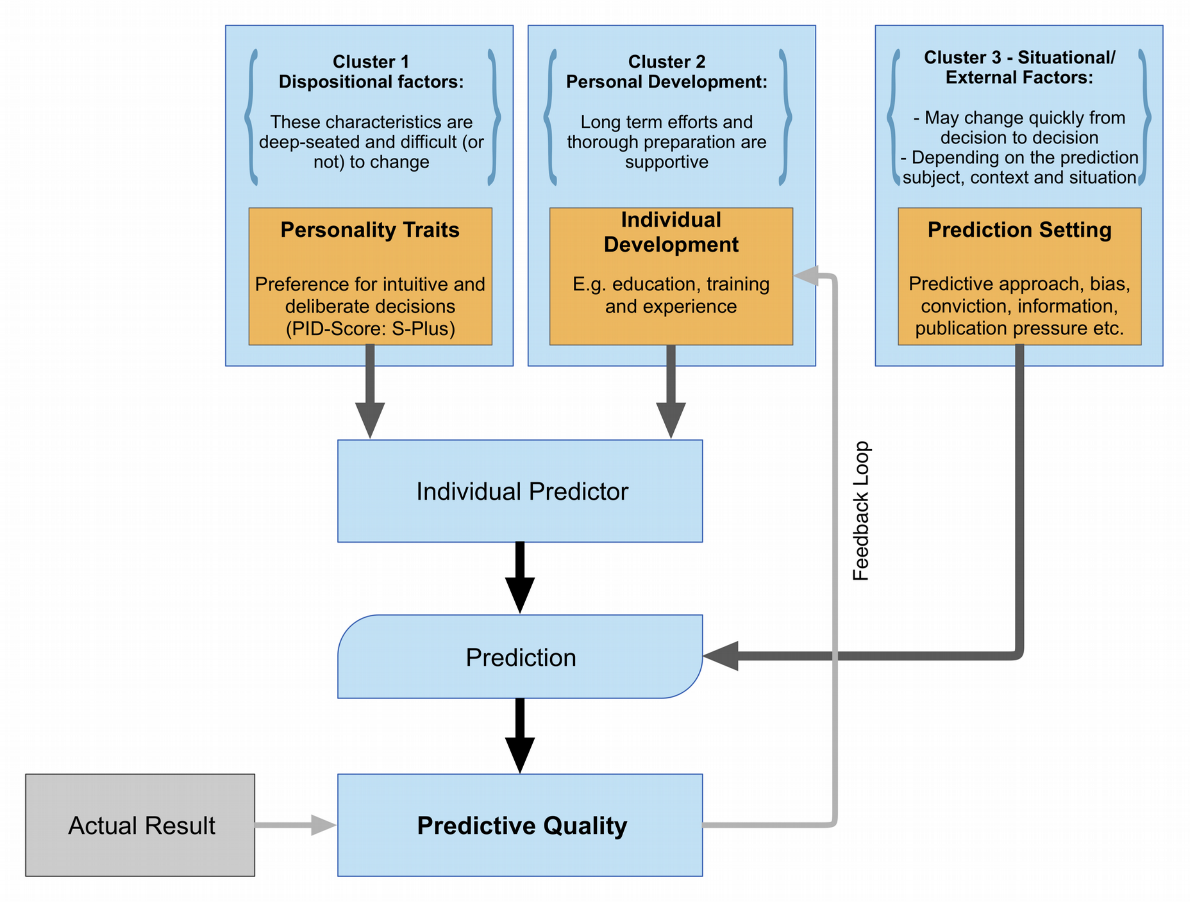


Figure 27: Deliberated Intuition Model

While there is no predictor who is always able to provide correct predictions there are predictors with superior performance. Additionally, the group design and group feedback loop in the experiment had only a minor impact on the predictive quality of the individuals. A pattern that was identified within the group of ‘superior predictors’ was that there are individual predictor characteristics, in particular the PID-S plus score, which indicate an individual preference for intuitive and deliberate decisions. However, as well as the personal predisposition of the predictor, there are more factors that contribute to good predictive quality. Superior predictions were observed from participants who used an approach which could be described as informed intuitive prediction. In-depth knowledge and training with deliberative forecasting methods, like formal prediction models and market experience, enhances predictive quality. The crux of the matter is that a top predictor combines these factors with their personality (e.g., the PID-S plus score), and their training and experience. Additionally, the setting of the actual prediction situation impacts their way of reasoning. The Deliberated Intuition Model can be regarded as a re-conceptualization and enhancement of dual process models, sometimes able to outperform the predictive quality of financial analysts using a traditional (rational) approach.

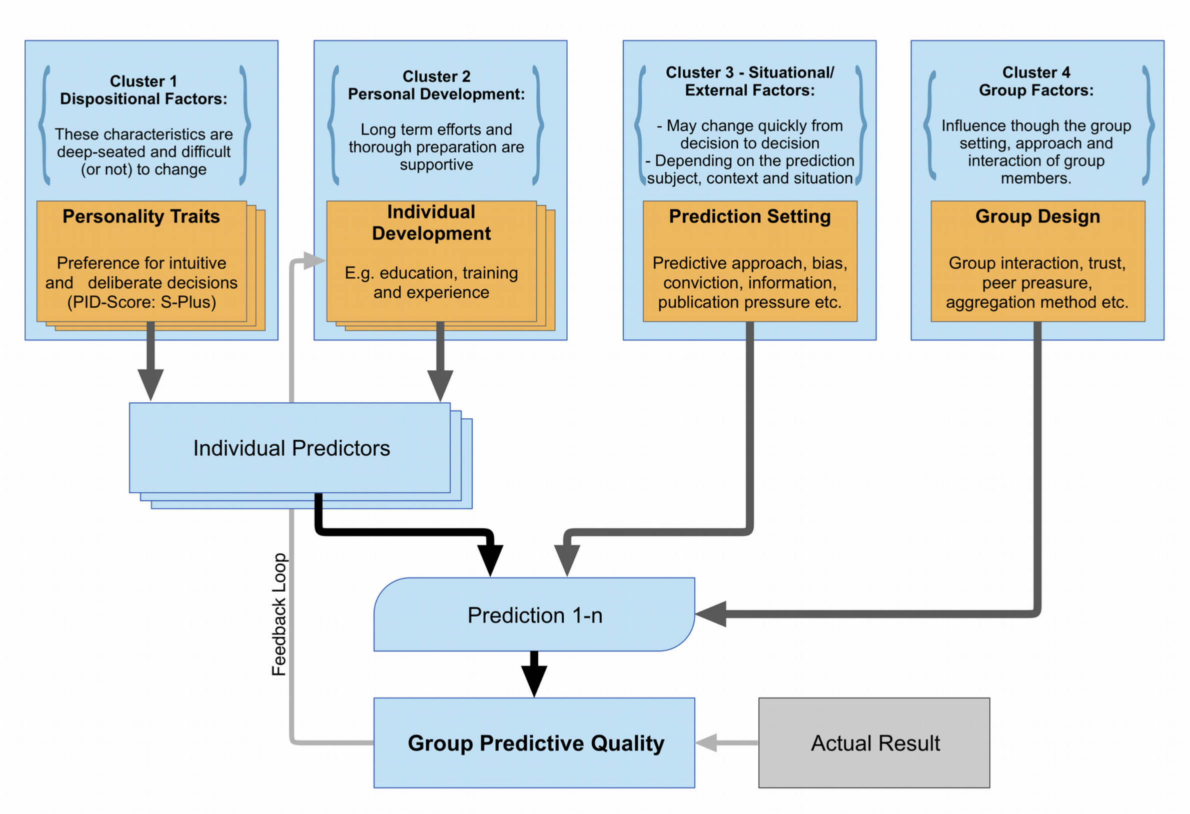


Figure 28: Deliberated Intuition Model for Groups

The application of this Deliberated Intuition Model has implications for the predictive quality of on-line or face-to-face groups making predictions. A careful group design with consideration of the mix of personality traits, the training and experience of the participants might improve the predictive quality of group output. Further research is needed to identify the optimal design, size, and composition of such groups. However, diversity in terms of intuition and (trained) deliberative perspectives within the group, and a process that facilitates both perspectives being considered in the group decision-making process might be the key to reaching superior predictive quality (see figure 28). The next step should be purposeful sampling to allow an in-depth assessment of the impact of the individual factors of figure 27 and 28, and their interdependencies in a group setting.

6.3 Synopsis and Conclusion

Group decision-making is a complex process (see also 2.3.2 Group decision-making). In order to accommodate this complexity, the mixed-methods approach in the form of a sequential study allowed us to address several issues. A purely positivist approach would have been appropriate for addressing questions 1 and 2, but to understand why this happens it was necessary to adopt a constructivist perspective as well. The mixed-methods approach was also suitable to address research question 3: what are the underlying key mechanisms, of the individual and of the group, that influence the decision-making process? The collective intelligence within the scope of investment decisions is still an exciting field. However, not all investment predictions by an online group are superior to financial analysts. A closer look at existing investment funds based on a collective intelligence approach revealed that—sometimes after an initial phase of outperformance—the performance might in some cases only be mediocre or worse. Although initial ideas about 'swarm intelligence' are already relatively old, there are still many open questions. This research project aimed to inform our understanding of the underlying processes and add to the body of knowledge. However, it is a truism that quite a few of the well-known research findings in relation to group decisions in business are implemented only partially or disregarded altogether in the business context. A central aspect of group decisions still seems to be the composition of the group and the specific questions asked. As a complement to existing analyst opinions these group decisions certainly have a high value, but the results are not necessarily superior. The question of the context in which group decisions are particularly good, and when rather problematic, is only known on some occasions, and many questions remain unanswered.

Despite these open questions, online communities in the investment sector seem increasingly to be finding their way. Quite a few industry insiders and commentators state that the financial industry is in a period of drastic changes (King, 2013; Skinner, 2014).

However, many of these promising approaches involve risks that have not been adequately studied and therefore could result in massive problems. In view of the potential benefits, however, it seems extraordinarily sensible from a business perspective for financial service providers to actively address these issues and to seek ways to integrate this external knowledge from the online communities into existing business processes or to establish new processes accordingly. There are several characteristics that may help to enhance group design. The Deliberated Intuition Model may help to prepare better group settings and improve predictive quality. Basically, digital social media seem to have the potential to significantly influence the business environment of financial services companies and to change them radically in some areas. Perhaps a practical approach is: *Groups can add value, but only with the appropriately selected individuals and setting*. If used correctly, social media, and in particular online groups, can digitally deliver a significant contribution to the value chain.

Areas for further research include improving the reliability and usability of the Deliberated Intuition Model, conducting an experiment on a larger scale, and experimenting with variations of group design and composition to identify those that are most effective. The Deliberated Intuition Model might also be useful to select members for select crowds (Mannes, Soll, & Larrick, 2014) or small crowds (Goldstein et al., 2014). Further research should also include variations in cultural context, investment instruments and market conditions, as well as the assessment of additional variables such as risk aversion, trading activity, over-optimism or sensation-seeking.

References

Acker, F. (2008). New findings on unconscious versus conscious thought in decision making: Additional empirical data and meta-analysis. *Judgment and Decision Making*, *3*(4), 292–303.

Aczel, B., Lukacs, B., Komlos, J., Aitken, M. R., & others. (2011). Unconscious intuition or conscious analysis? Critical questions for the Deliberation-Without-Attention paradigm. *Judgment and Decision Making*, *6*(4), 351–358.

Aiolfi, M., Rodriguez, M., & Timmermann, A. (2009). Understanding Analysts' Earnings Expectations: Biases, Nonlinearities, and Predictability. *Journal of Financial Econometrics*, *8*(3), 305–334. https://doi.org/10.1093/jjfinec/nbp024

Akerlof, G. (2009). *Animal spirits : How human psychology drives the economy, and why it matters for global capitalism*. Princeton: Princeton University Press.

Alter, S. L. (1980). *Decision support systems : Current practice and continuing challenges*. Reading Mass. [etc.]: Addison-Wesley Publ.

Anderson, G. L., & Herr, K. (1999). The new paradigm wars: Is there room for rigorous practitioner knowledge in schools and universities? *Educational Researcher*, *28*(5), 12–21.

Anderson, G. L., & Herr, K. (2005). *The action research dissertation : a guide for students and faculty*. Thousand Oaks: SAGE Publications.

Aramburo, S. A., Acevedo, J. A. C., & Morales, Y. O. (2009). System Dynamics and Laboratory Experiments (pp. 78–93). Presented at the 27th International Conference of the System Dynamics Society 2009, Albuquerque, New Mexico, USA. Retrieved from http://www.systemdynamics.org/conferences/2009/proceed/papers/P1123.pdf

Ariely, D. (2009). *Predictably irrational : The hidden forces that shape our decisions*. London: Harper.

Ariely, D. (2010). *The upside of irrationality : The unexpected benefits of defying logic at work and at home* (1st ed.). New York: Harper.

Armstrong, J. S. (2001). Combining Forecasts. In J. S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners* (pp. 417–439). Norwell, MA: Kluwer/Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-0-306-47630-3_19

Arnott, D. D., & Pervan, G. G. (2005). A critical analysis of decision support systems research. *Journal of Information Technology*, *20*(2), 67–87.

Asch conformity experiments. (n.d.). Retrieved 29 December 2010, from http://en.academic.ru/dic.nsf/enwiki/343898

Asch, S. E. (1952). *Social psychology.* Englewood Cliffs: Prentice-Hall.

Asch, S. E. (1956). *Studies of independence and conformity : I: A minority of one against a unanimous majority*. Washington: American Psychological Ass.

Bachelier, L. (1900). *Théorie de la spéculation [Theory of Speculation]*. Gauthier-Villars, Paris.

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2006). *Multivariate Analysemethoden: eine anwendungsorientierte Einführung [Multivariate analysis: an application oriented introduction]* (11th ed.). Berlin: Springer.

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (Eds.). (2016). *Multivariate Analysemethoden: eine anwendungsorientierte Einführung [Multivariate analysis: an application oriented introduction]* (14. Ed.). Berlin: Springer Gabler.

Ball, S., & Wingender, J. R. (1988). A Simulation of Investment Analysis, Portfolio Management and Reporting Using LOTUS 1-2-3. *Developments in Business Simulation and Experiential Learning*, *15*. Retrieved from https://journals.tdl.org/absel/index.php/absel/article/view/1917

Ballantyne, R., Hughes, K., & Bond, N. (2016). Using a Delphi approach to identify managers' preferences for visitor interpretation at Canterbury Cathedral World Heritage Site. *Tourism Management*, *54*, 72–80. https://doi.org/10.1016/j.tourman.2015.10.014

Barberis, N., & Thaler, R. (2003). Chapter 18 A survey of behavioral finance. In *Handbook of the Economics of Finance* (Vol. Volume 1, Part B, pp. 1053–1128). Elsevier. https://doi.org/10.1016/S1574-0102(03)01027-6

Bartunek, J., & Louis, M. R. (1996). *Insider/outsider team research*. Thousand Oaks: SAGE Publications.

Batchelor, R., & Dua, P. (1992). Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting*, *11*(2), 169–181. https://doi.org/10.1002/for.3980110207

Baumeister, R., & Leary, M. (1997). Writing narrative literature reviews. *Review of General Psychology*, *1*(3), 311–320.

Beechey, M., Gruen, D., & Vickery, J. (2000). *The efficient market hypothesis: A survey*. Citeseer.

Berg, J. E., & Rietz, T. A. (2003). Prediction markets as decision support systems. *Information Systems Frontiers*, *5*(1), 79–93.

Berg, J. E., & Rietz, T. A. (2006). The Iowa Electronic Markets: Stylized facts and open issues.

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. https://doi.org/10.1024/0170-1789.25.4.179

Beyer, A., & Guttman, I. (2011). The Effect of Trading Volume on Analysts' Forecast Bias. *The Accounting Review*, *86*(2), 451–481. https://doi.org/10.2308/accr.00000030

BigCharts - Interactive Charting. (2015, December 26). Retrieved 26 December 2015, from http://bigcharts.marketwatch.com/advchart/frames/frames.asp?show=&insttype=Fund&symb=MOFQX&time=20&startdate=1%2F4%2F1999&enddate=12%2F26%2F2015&freq=2&compidx=SP500&comptemptext=&comp=none&ma=0&maval=9&uf=0&lf=1&lf2=0&lf3=0&type=2&style=320&size=2&x=48&y=4&timeFrameToggle=false&compareToToggle=false&indicatorsToggle=false&chartStyleToggle=false&state=12

Bollen, J., Mao, H., & Zeng, X. J. (2010). Twitter mood predicts the stock market. *Arxiv Preprint arXiv:1010.3003*.

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

Bolton, G. E. J. (2010). *Reflective practice writing and professional development.* London: SAGE Publications.

Bordens, K. S., & Horowitz, I. A. (2001). *Social psychology* (2nd ed.). Mahwah N.J.: Lawrence Erlbaum Associates.

Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler [Statistics for Human and Social Scientists]*. Berlin; Heidelberg; New York: Springer.

Bortz, J., & Döring, N. (2015). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler [Research Methods and Evaluation: for Human and Social Scientists]* (4th edition, limited special edition). Berlin: Springer.

Bradbury, H., & Reason, P. (2001). Conclusion: Broadening the Bandwidth of Validity: Issues and choice points for improving quality in action research. In B. Bradbury & P. Reason (Eds.), *Handbook of Action Research: Participative inquiry and practice* (pp. 447–455). London: SAGE Publications.

Bradbury, M. A., Hens, T., & Zeisberger, S. (2014). Improving investment decisions with simulated experience. *Review of Finance*, rfu021.

Bradshaw, M. T., Richardson, S. A., & Sloan, R. G. (2006). The relation between corporate financing activities, analysts' forecasts and stock returns. *Journal of Accounting and Economics*, *42*(1), 53–85.

Brafman, O., & Brafman, R. (2009). *Sway : The irresistible pull of irrational behaviour*. London: Virgin.

Brahm, C., & Kleiner, B. H. (1996). Advantages and disadvantages of group decision-making approaches. *Team Performance Management*, *2*(1), 30.

Brauer, M., Judd, C. M., & Gliner, M. D. (1995). The effects of repeated expressions on attitude polarization during group discussions. *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY*, *68*(6), 1014.

Bredenbals, J. (2015, September 28). Markt & Strategie - FONDS professionell Multimedia GmbH. Retrieved 26 December 2015, from http://www.fondsprofessionell.de/news/markt-strategie/nid/schwarmintelligenz-fonds-von-hauck-amp-aufhaeuser-wird-geschlossen/gid/1022899/ref/1/

Bryman, A., & Bell, E. (2003). *Business research methods*. Oxford, New York: Oxford University Press.

Bullough, R. V., & Pinnegar, S. (2001). Guidelines for Quality in Autobiographical Forms of Self-Study Research. *Educational Researcher*, *30*(3), 13–21.

Camerer, C. F., Loewenstein, G., & Rabin, M. (2011). *Advances in Behavioral Economics*. Princeton University Press. Retrieved from https://books.google.de/books?id=sA4jJOjwCW4C

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental design for research*. Chicago: Rand McNally.

Chase, R. (2015). *Peers Inc: how people and platforms are inventing the collaborative economy and reinventing capitalism*. London: Headline Publ.

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.

Chou, C. (2002). Developing the E-Delphi system: A web-based forecasting tool for educational research. *British Journal of Educational Technology*, *33*(2), 233–36.

Chung, J., & Geoffrey, Q. P. (n.d.). Introduction to GDSS. Retrieved 24 October 2010, from http://home.bre.polyu.edu.hk/~bsjacky/gdss/L4a.htm#anchor3

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

Codol, J.-P. (1975). 'Effet PIP' et conflit de normes ['PIP Effect' and conflict of standards].

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

Combs, J. P., & Onwuegbuzie, A. J. (2010). Emergent data analysis techniques in mixed methods research: A synthesis. In A. Tashakkori & C. Teddli (Eds.), *SAGE Handbook of Mixed Methods in Social Behavioral Research* (2nd ed., pp. 397–430). Los Angeles: SAGE Publications.

Condorcet, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix [Essay on the application of analysis to the probability of decisions made by a majority of votes]*. Retrieved from http://books.google.de/books?id=RzAVAAAAQAAJ

Connelly, F. M., & Clandinin, D. J. (1990). Stories of Experience and Narrative Inquiry. *Educational Researcher*, *19*, 2–14. https://doi.org/10.3102/0013189X019005002

Cooper, D. J., & Kagel, J. H. (2005). Are two heads better than one? Team versus individual play in signaling games. *American Economic Review*, *95*(3), 477–509.

Cooper, D., & Schindler, P. S. (2011). *Business research methods* (11th ed.). New York: McGraw-Hill Higher Education.

Cornett, M. M., Mehran, H., & Tehranian, H. (1998). Are Financial Markets Overly Optimistic about the Prospects of Firms That Issue Equity? Evidence from Voluntary versus Involuntary Equity Issuances by Banks. *The Journal of Finance*, *53*(6), 2139–2159. https://doi.org/10.1111/0022-1082.00085

Cornwall, A. (1996). Towards participatory practice: participatory rural appraisal (PRA) and the participatory process. In K. De Koning & M. Martin (Eds.), *Participatory Research in Health: Issues and Experiences* (pp. 94–107). London: Zed Books.

Courtault, J. M., Kabanov, Y., Bru, B., Crépel, P., Lebon, I., & Le Marchand, A. (2000). Louis Bachelier on the Centenary of Théorie de la Spéculation. *Mathematical Finance*, *10*(3), 339–353.

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

Cuhls, K. (n.d.). Delphi method. Fraunhofer Institute for Systems and Innovation Research, Germany. Retrieved from http://www.unido.org/fileadmin/import/16959_DelphiMethod.pdf

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. (n.d.). The Delphi methodology. Retrieved 30 March 2011, from http://www.fernuni-hagen.de/ZIFF/v2-ch45a.htm

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.

Damodaran, A. (2006). *Damodaran on valuation : Security analysis for investment and corporate finance* (2nd ed.). Hoboken: John Wiley & Sons.

Davis, M., & Etheridge, A. (2006). *Louis Bachelier's theory of speculation : The origins of modern finance*. Princeton; Oxford: Princeton University Press.

Davison, R. M. (2001). A survey of group support systems: Technology and operation. *Sprouts: Working Papers on Information Systems*, *1*(12). Retrieved from http://sprouts.aisnet.org/1-12

De Bondt, W. F. M., & Thaler, R. H. (2002). Do Analysts Overreact? In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: the psychology of intuitive judgement*. New York: Cambridge University Press.

de Vries, M., Holland, R. W., & Witteman, C. L. M. (2008). Fitting decisions: Mood and intuitive versus deliberative decision strategies. *Cognition and Emotion*, *22*(5), 931–943. https://doi.org/10.1080/02699930701552580

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., … Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 201517441. https://doi.org/10.1073/pnas.1517441113

DeSanctis, G., & Gallupe, B. (1984). *Information system support for group decision making*. Minneapolis: Management Information Systems Research Center, Graduate School of Management, University of Minnesota.

DeSanctis, G., & Gallupe, B. (1987). A foundation for the study of group decision support systems. *Management Science, 33*(5), 589–609.

Deutsch, M., & Krauss, R. M. (1965). *Theories in social psychology*. New York: Basic Books.

Dhami, S. (2016). *The Foundations of Behavioral Economic Analysis*. OUP Oxford. Retrieved from https://books.google.de/books?id=NSahDQAAQBAJ

Diehl, M., & Stroebe, W. (1990). *Productivity loss in idea-generating groups : Tracking down the blocking-effect*. Tübingen: Univ.

Duckworth, W. E., Gear, A. E., & Lockett, A. G. (1977). *A guide to operational research*. London; New York: Chapman and Hall ; Wiley.

Dueck, G. (2015). *Schwarmdumm: so blöd sind wir nur gemeinsam [Swarm Stupidity: We Can Only Be this Stupid if We All Work Together]*. Frankfurt New York: Campus Verlag.

Dupernex, S. (2007). Why might share prices follow a random walk? *Student Economic Review*, (21).

Eckl-Dorna, W. (2013, April 18). Energieversorger: RWE stimmt Aktionäre auf Durststrecke ein [Utilities: RWE agrees shareholders on a losing streak]. Retrieved 23 April 2014, from http://www.manager-magazin.de/unternehmen/energie/a-895176.html

Edwards, R. D., & Magee, J. (1997). *Technical analysis of stock trends*. Chicago; New York: J. Magee ; AMACOM.

Ellis, A. K. (1977). *Teaching and learning elementary social studies*. Boston: Allyn and Bacon.

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

Endress, T. (2013). A Pilot Study of Quality of Equity Predictions in Online Communities. *International Journal of Innovations in Business*, *3*(2), 225–240.

Endress, T. (2014). Forecasting Behaviour of Financial Analysts in Open Competition and in Private. In D. Vrontis, Y. Weber, & E. Tsoukatos (Eds.), *Proceedings of 7th Annual Conference of the EuroMed Academy of Business: The Future of Entrepreneurship* (pp. 519–527). Kristiansand, Norway: EuroMed Press. Retrieved from http://media.wix.com/ugd/390e49_6e4e3f4324af43e6a660ebd9d1b0ce62.pdf

Endress, T. (2015). An e-Delphi experiment of quality of equity predictions in online groups. *Qualitative Research in Financial Markets*, (7). https://doi.org/10.1108/QRFM-05-2014-0014

Endress, T., & Gear, T. (2013a). e-Delphi Pilot Experiment of Quality of Equity Predictions in Online Groups. *International Journal of Management Cases*, *15*(4), 74–88.

Endress, T., & Gear, T. (2013b). e-Delphi Pilot Experiment of Quality of Equity Predictions in Online-Groups. In G. Vignali (Ed.), *Instituto Politécnico (Viana do Castelo)* (p. 58). Lancashire: Access Press.

Endress, T., & Gear, T. (2015). Stock Prices: Are Intuitive or Deliberate Persons Better Forecasters? *ECONOMICS & SOCIOLOGY*, *8*(4), 43–50. https://doi.org/10.14254/2071-789X.2015/8-4/3

Eom, S. B. (2001). Decision Support Systems. In *International Encyclopedia of Business and Management* (2nd Edition). London: International Thomson Business Publishing Co. Retrieved from http://e-learning.dmst.aueb.gr/mis/Cases/Nederlandse_Spoorwegen/Case/Training_Files/iebmdssrwweb.pdf

Eriksson, P., & Kovalainen, A. (2008). *Qualitative methods in business research*. London: SAGE Publications.

Everling, O., & Lempka, R. (Eds.). (2016). *Finanzdienstleister der nächsten Generation: Megatrend Digitalisierung: Strategien und Geschäftsmodelle [Next Generation Financial Services: Megatrend Digitization: Strategies and Business Models]*. Frankfurt am Main: Frankfurt School Verlag.

Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business*, *38*(1), 34–105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, *25*(2), 383–417.

Fellner-Röhling, G., & Maciejovsky, B. (2007). Risk attitude and market behavior: Evidence from experimental asset markets. *Journal of Economic Psychology*, *28*(3), 338–350.

Finlay, L. (2002). 'Outing' the researcher: The provenance, process, and practice of reflexivity. *Qualitative Health Research*, *12*(4), 531–545.

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

Fisher, I. (1930). *The theory of interest as determined by impatience to spend income and opportunity to invest it*. New York: The Macmillan Company.

Fisher, L. (2009). *The perfect swarm : The science of complexity in everyday life*. New York: Basic Books.

Fjermestad, J., & Hiltz, S. R. (1998). An assessment of group support systems experimental research: Methodology and results. *J. Manage. Inf. Syst.*, *15*(3), 7–149.

Fjerrnestad, J., & Hiltz, S. R. (2002). Experimental studies of group decision support systems: An assessment of variables studied and methodology. In *System Sciences, 1997, Proceedings of the Thirtieth Hawaii International Conference on* (Vol. 2, pp. 45–65).

Fleischer, J. (2005). *Die Performance von Analystenempfehlungen [The performance of analyst recommendations]* (1. Aufl.). Wiesbaden: Gabler.

Flynn, R. R., & Druzdzel, M. J. (2003). Decision Support Systems. In A. Kent (Ed.), *Encyclopedia of Library and Information Science*. New York: Marcel Dekker. Retrieved from http://www.pitt.edu/~druzdzel/psfiles/dss.pdf

Flyvbjerg, B. (2011). What Is a Case study? In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research* (4th Edition, pp. 301–316). London; Thousand Oaks: SAGE Publications. Retrieved from http://www.sbs.ox.ac.uk/centres/bt/directory/Documents/CaseStudy4%202HBQR11PRINT.pdf

Forsyth, D. R. (1990). *Group dynamics*. Pacific Grove: Brooks/Cole Pub. Co.

Fox, J. (2009). *The myth of the rational market : A history of risk, reward, and delusion on Wall Street* (1st ed.). New York: Harper Business.

Freifeld, J. (1996). Speculative Bubbles: Financial genius before the fall. Retrieved 29 December 2010, from http://www.clarity.net/~jake/bubble.htm

Galton, F. (1907, March 7). Vox Populi. *Nature*, *75*, 450–451.

Garvin, D. A. (2003). Making the case. *Harvard Magazine*, *106*(1), 56–65.

Gear, A. E., Marsh, N. R., & Sergent, P. (1985). Semi-Automated feedback and team behavior. *Human Relations*, *38*(8), 707–721.

Gear, A. E., & Read, M. J. (1993). On-line group process support. *OMEGA -OXFORD-PERGAMON PRESS-*, *21*(3), 261.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge : the dynamics of science and research in contemporary societies*. London; Thousand Oaks; New Delhi: SAGE Publications.

Gigerenzer, G. (2008). *Bauchentscheidungen : Die Intelligenz des Unbewussten und die Macht der Intuition [Gut Feelings: The Intelligence of the Unconscious]* (1. Aufl., Taschenbuchausg.). München: Goldmann.

Gigone, D., & Hastie, R. (1993). The common knowledge Effect: Information sharing and group judgment. *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY*, *65*(5), 959.

Gilovich, T., Griffin, D. W., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: the psychology of intuitive judgement*. Cambridge, U.K. ; New York: Cambridge University Press.

Glöckner, A., & Witteman, C. (2010). Beyond dual-process models: A categorisation of processes underlying intuitive judgement and decision making. *Thinking & Reasoning*, *16*(1), 1–25. https://doi.org/10.1080/13546780903395748

Goedhart, M., Raj, R., & Saxena, A. (2010, April). Equity analysts: Still too bullish. *Mckinseyquarterly.com*. Retrieved from http://www.mckinsey.com/insights/corporate_finance/equity_analysts_still_too_bullish

Goldstein, D. G., McAfee, R. P., & Suri, S. (2014). The wisdom of smaller, smarter crowds (pp. 471–488). ACM Press. https://doi.org/10.1145/2600057.2602886

Gordon, M. J. (1962). *The investment, financing, and valuation of the corporation.* Homewood, Ill.: R.D. Irwin.

Grabill, S. J. (Ed.). (2007). *Sourcebook in late-scholastic monetary theory : The contributions of Martín de Azpilcueta, Luis de Molina, S.J., and Juan de Mariana, S.J.* Lanham MD: Lexington Books.

Graham, B. (2003). *The intelligent investor* (Rev. ed.). New York: HarperBusiness Essentials.

Gray, L., & Diehl, P. (1992). *Research methods for business and management*. New York [etc.]: Macmillan.

Greenwood, D. J. (2007). Pragmatic Action Research. *International Journal of Action Research*, *3*(1+ 2), 131–148.

Grinblatt, M., & Keloharju, M. (2009). Sensation Seeking, Overconfidence, and Trading Activity. *The Journal of Finance*, *64*(2), 549–578. https://doi.org/10.1111/j.1540-6261.2009.01443.x

Gruenfeld, D., Mannix, E., Williams, K., & Neale, M. (1996). Group composition and decision making: How member familiarity and information distribution affect process and performance. *Organizational Behavior and Human Decision Processes*, *67*(1), 1.

Habermas, J. (1971). *Knowledge and human interests*. Boston: Beacon Press.

Harteis, C., & Gruber, H. (2008). Intuition and Professional Competence: Intuitive Versus Rational Forecasting of the Stock Market. *Vocations and Learning*, *1*(1), 71–85. https://doi.org/10.1007/s12186-007-9000-z

Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review*, *35*(4).

Heron, J. (1996). *Co-operative inquiry : research into the human condition.* London: SAGE Publications.

Hessels, L. K., & van Lente, H. (2008). Re-thinking new knowledge production: A literature review and a research agenda. *Research Policy*, *37*(4), 740–760.

Ho, T. Q. (2012). *Three Essays on Financial Analysts' Stock Price Forecasts*. University of Manchester, Manchester. Retrieved from https://www.escholar.manchester.ac.uk/api/datastream?publicationPid=uk-ac-man-scw:186308&datastreamId=FULL-TEXT.PDF

Hofstadter, D. R. (1979). *Gödel, Escher, Bach : An eternal golden braid*. New York: Basic Books.

Hogg, M. A., & Cooper, J. (2003). *The Sage handbook of social psychology*. London; Thousand Oaks: SAGE Publications.

Hogg, M. A., Turner, J., & Davidson, B. (1990). Polarized norms and social frames of reference: A test of the self-categorization theory of group polarization. *Basic and Applied Social Psychology*, *11*(1), 77–100.

Höhler, G. (2013, April 14). Korruptionsermittlungen: Griechische Richter gegen Siemens-Manager | ZEIT ONLINE [Corruption investigations: Greek judges against Siemens manager | ZEIT ONLINE]. Retrieved 23 April 2014, from http://www.zeit.de/wirtschaft/2013-04/siemens-korruption-giechenland-von-pierer

Howe, J. (2008). *Crowdsourcing : How the power of the crowd is driving the future of business*. London: Random House Business.

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

Huber, G. P. (1984). Issues in the design of group decision support sytems. *MIS Quarterly*, *8*(3).

Hui, K. W., Wei, J., & You, H. (2013). *Analyst Incentives, Forecast Biases, and Stock Returns*. Hong Kong University of Science and Technology. Retrieved from http://sa.shufe.edu.cn/upload/_info/142423_1312170220581.pdf

Intelligent Recommendations GmbH. (2015). Intelligent Recommendations Global Growth Fund. Retrieved 26 December 2015, from http://www.ir-system.com/fonds/

INVESTTOR. (2015). Retrieved 26 December 2015, from https://www.investtor.de/

Janis, I. L. (1982). *Groupthink : Psychological studies of policy decisions and fiascoes* (2nd ed.). Boston: Houghton Mifflin.

Janssen, J., & Laatz, W. (2007). Statistische Datenanalyse mit SPSS für Windows: Eine anwendungsorientierte Einführung [Statistical Data Analysis with SPSS for Windows: An application-oriented introduction]. Retrieved from http://dx.doi.org/10.1007/978-3-540-72978-5

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

Jose, V. R. R., Grushka-Cockayne, Y., & Jr, K. C. L. (2014). Trimmed Opinion Pools and the Crowd's Calibration Problem. *Management Science*, *60*(2), 463–475.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, *64*(6), 515–526. https://doi.org/10.1037/a0016755

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Kahneman, D., & Tversky, A. (2000). *Choices, values, and frames*. New York; Cambridge: Russell sage Foundation; Cambridge University Press.

Kaplan, C. A. (2001). Collective intelligence: A new approach to stock price forecasting. In *IEEE INTERNATIONAL CONFERENCE ON SYSTEMS MAN AND CYBERNETICS* (Vol. 5, pp. 2893–2898).

Keen, P. G. W., & Morton, M. S. (1978). *Decision support systems : An organizational perspective*. Reading: Addison-Wesley Pub. Co.

Keller, C., & Siegrist, M. (2006). Investing in stocks: The influence of financial risk attitude and values-related money and stock market attitudes. *Journal of Economic Psychology*, *27*(2), 285–303. https://doi.org/10.1016/j.joep.2005.07.002

Kellerman, B. (2014, November 19). Forecast bias: Are Australian analysts overly optimistic? Retrieved 9 January 2016, from https://finsia.com/news/news-article/2014/11/18/forecast-bias-are-australian-analysts-overly-optimistic

Keynes, J. M. (1936). *The general theory of employment, interest and money*. New York: Harcourt, Brace & World.

King, B. (2013). *Bank 3.0: why banking is no longer somewhere you go, but something you do*. Singapore: Marshall Cavendish Business.

Koller, T., & McKinsey and Company. (2010). *Valuation : Measuring and managing the value of companies* (5th ed.). Hoboken: Wiley.

Komáromi, G. (2006). *Anatomy of stock market bubbles*. Nagarjuna Hills: ICFAI UP.

Kuckartz, U. (2014). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung [Qualitative Text Analysis: A Guide to Methods, Practice and Using Software]*. Weinheim, Bergstr: Beltz Juventa.

Kuhn, J. (2004). *Kommerzielle Nutzung mobiler Anwendungen-Ergebnisse der Delphi-Studie ‘Mobile Business’ [Commercial use of mobile applications results of the Delphi study ‘Mobile Business.’]*. Retrieved from http://epub.uni-regensburg.de/10163/1/Kommerzielle_Nutzung_mobiler_Anwendungen.pdf

Larson, J. R., Foster-Fishman, P. G., & Keys, C. B. (1994). Discussion of shared and unshared information in decision-making groups. *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY*, *67*(3), 446.

Le Bon, G. (2009). *Psychologie der Massen [Psychologie des foules]*. Hamburg: Nikol.

Leamer, E. (2003). Bubble trouble. *Your Home Has a P/E Ratio Too, UCLA Anderson Forecast Report, June*.

Legerstee, R., & Franses, P. H. (2015). Does Disagreement Amongst Forecasters Have Predictive Value? *Journal of Forecasting*, *34*(4), 290–302. https://doi.org/10.1002/for.2330

Lewin, K. (1946). Action Research and Minority Problems. *Journal of Social Issues*, *2*(4), 34–46.

Lichtendahl, K. C., Grushka-Cockayne, Y., & Pfeifer, P. E. (2013). The Wisdom of Competitive Crowds. *Operations Research*, *61*(6), 1383–1398.

Lin, H., & McNichols, M. F. (1998). Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics*, *25*(1), 101–127. https://doi.org/10.1016/S0165-4101(98)00016-0

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, Calif.: Sage Publications.

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

Liu, W.-N., & Yao, J. (n.d.). Web-based dynamic delphi: A new survey instrument. Retrieved from http://www2.cs.uregina.ca/~jtyao/Papers/Delphi_Cam.pdf

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

Malmendier, U., & Shanthikumar, D. (2007). *Do Security Analysts Speak in Two Tongues?* (No. w13124). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w13124.pdf

Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology*, *107*(2), 276–299. https://doi.org/10.1037/a0036677

Marinovic, I., & Ottaviani, M. (2013). Forecasters' Objectives and Strategies. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2B). Amsterdam ; Boston: Elsevier North-Holland.

Mauboussin, M. J. (2007). Explaining the wisdom of crowds. Retrieved from http://www.lmcm.com/pdf/Explaining_the_Wisdom_of_Crowds.pdf

Mayring, P. (2002). *Einführung in die qualitative Sozialforschung : eine Anleitung zu qualitativem Denken [Introduction to qualitative research: a guide to qualitative thinking]*. Weinheim: Beltz.

Mayring, P. (2010). *Qualitative Inhaltsanalyse Grundlagen und Techniken [Qualitative content analysis principles and techniques]* (11th ed.). Weinheim: Beltz.

McNiff, J., & Whitehead, J. (2009). *Doing and writing action research*. London: SAGE Publications.

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

Michel, J.-S., & Pandes, J. A. (2012, March 15). *Are Analysts Really Too Optimistic?* Retrieved from http://www.mfsociety.org/modules/modDashboard/uploadFiles/conferences/MC19~508~p16sub1m3ll90h4l11koavm12mk4.pdf

Milgram, S. (1964). Group pressure and action against a person. *Journal of Abnormal and Social Psychology*, *69*(2), 137–143. https://doi.org/10.1037/h0047759

Miller, P. (2010). *The smart swarm : How understanding flocks, schools, and colonies can make us better at communicating, decision making, and getting things done*. New York: Avery.

Miller, Z. (2010). *Tradestream your way to profits : Building a killer portfolio in the age of social media*. Hoboken: Wiley.

Morningstar, Inc. (2008). The Morningstar Rating for funds. Retrieved from http://corporate.morningstar.com/US/documents/MethodologyDocuments/FactSheets/MorningstarRatingForFunds_FactSheet.pdf

Morningstar, Inc. (2010, March 11). MOFQX Marketocracy Masters 100, mutual funds, quote, price - Morningstar. Retrieved 4 November 2010, from http://quote.morningstar.com/fund/f.aspx?t=MOFQX

Mullen, B., Johnson, C., & Salas, E. (1991). Productivity loss in brainstorming groups: A meta-analytic integration. *Basic and Applied Social Psychology*, *12*(1), 3–23.

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

Nijstad, B. A. (2009). *Group performance*. Hove; New York: Psychology Press.

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (2003). Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology.*, *39*(6), 531.

Nunamaker, J. F. (1997). Future research in group support systems: Needs, some questions and possible directions. *Int. J. Hum.-Comput. Stud.*, *47*(3), 357–385.

O'Brien, R. (1998). Overview of Action Research Methodology [Faculty of Information Studies, University of Toronto]. Retrieved 28 August 2011, from http://web.ca/~robrien/papers/arfinal.html

O'Malley, M. (2010). *The wisdom of bees : what the hive can teach business about leadership, efficiency, and growth*. New York: Portfolio.

Osband, K. (1989). Optimal Forecasting Incentives. *Journal of Political Economy*, *97*(5), 1091–1112.

Osborn, A. F. (1963). *Applied imagination : Principles and procedures of creative problem-solving. 3rd rev ed.* New York: Charles Scribners Sons.

Ott, K. (2012, December 5). Korruption bei Thyssen-Krupp - Ärger mit dem Schienenpapst [Corruption at ThyssenKrupp - trouble with the rail Pope]. Retrieved 23 April 2014, from http://www.sueddeutsche.de/wirtschaft/korruption-bei-thyssen-krupp-aerger-mit-dem-schienenpapst-1.1541749

Ottaviani, M., & Sørensen, P. N. (2006). The strategy of professional forecasting. *Journal of Financial Economics*, *81*(2), 441–466. https://doi.org/10.1016/j.jfineco.2005.08.002

Page, S. E. (2008a). An introduction to modeling: Scott E Page. Retrieved 16 January 2011, from http://www.cscs.umich.edu/~spage/teaching_files/modeling_lectures/MODEL5/M18predicmodel.htm

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

Patton, M. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park: Sage Publications.

Paulus, P. B., Dzindolet, M. T., Poletes, G., & Camacho, L. M. (1993). Perception of performance in group brainstorming: The illusion of group productivity. *PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN*, *19*(1), 78.

Penman, S. H. (1998). A synthesis of equity valuation techniques and the terminal value calculation for the dividend discount model. *Review of Accounting Studies*, *2*(4), 303–323.

Penman, S. H. (2007). *Financial statement analysis and security valuation*. London: McGraw-Hill.

Pezaro, S., & Clyne, W. (2015). Achieving Consensus in the Development of an Online Intervention Designed to Effectively Support Midwives in Work-Related Psychological Distress: Protocol for a Delphi Study. *JMIR Research Protocols*, *4*(3), e107. https://doi.org/10.2196/resprot.4766

Piros, C. D., & Pinto, J. E. (Eds.). (2013). *Economics for investment decision makers: micro, macro, and international economics*. Hoboken, N.J: Wiley.

Plano, C. V. L., & Badiee, M. (2010). Research questions in mixed methods research. In A. Tashakkori & C. Teddli (Eds.), *SAGE handbook of mixed methods in social behavioral research* (2nd ed., pp. 275–304). Los Angeles: SAGE Publications.

Power, D. J. (2002). *Decision support systems : Concepts and resources for managers*. Westport, Conn.: Quorum Books.

Preacher, K. J. (2001). *Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence*. Retrieved from http://quantpsy.org/chisq/chisq.htm

Price, M. C., & Norman, E. (2008). Intuitive decisions on the fringes of consciousness: Are they conscious and does it matter? *Judgment and Decision Making*, *3*(1), 28.

Quattrociocchi, W., Caldarelli, G., & Scala, A. (2014). Opinion dynamics on interacting networks: media competition and social influence. *Scientific Reports, 4*. https://doi.org/10.1038/srep04938

Randolph, J. (2009). A guide to writing the dissertation literature review. *Practical Assessment, Research & Evaluation*, *14*(13). Retrieved from http://pareonline.net/pdf/v14n13.pdf

Read, M. J., & Gear, A. E. (1994). Teamworking to develop technology strategy. *INTERNATIONAL JOURNAL OF TECHNOLOGY MANAGEMENT*, *9*(2), 241.

Ricardo, D. (1817). *On the principles of political economy and taxation*. London: John Murray.

Rietz, T. A., Forsythe, R., & Berg, J. E. (1997). The Iowa Electronic Market. In D. Paxson & D. Wood (Eds.), *Blackwell Encyclopedic Dictionary of Finance* (pp. 111–113). Oxford: Blackwell.

Rohrbach, B. (1969). Kreativ nach Regeln [Creative according to rules]. *Absatzwirtschaft, 19*(12), 73–76.

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. https://doi.org/10.1016/j.techfore.2004.03.004

Rubenfeld, S. (2013, January 16). Korruption: Whistleblower verklagt Siemens [Corruption Whistleblower sued Siemens]. Retrieved 23 April 2014, from http://www.spiegel.de/wirtschaft/unternehmen/korruption-whistleblower-verklagt-siemens-a-877995.html

Ryan, B. (2007). *Corporate finance and valuation*. London: Thomson Learning.

Ryan, B., Scapens, R. W., & Theobald, M. (2002). *Research method and methodology in finance and accounting*. London: Thomson.

Saavedra, E. (1996). Teachers study groups: Contexts for transformative learning and action. *Theory Into Practice*, *35*(4), 271–277. https://doi.org/10.1080/00405849609543734

Sample, J. A. (1984, March). Nominal Group Technique: An alternative to brainstorming. *Journal of Extension*, *22*(2). Retrieved from http://www.joe.org/joe/1984march/iw2.php

Sauter, V. L. (2001, March 25). GDSS. Retrieved 12 December 2010, from http://www.umsl.edu/~sauterv/DSS/gdss_defined.html

Schön, D. A. (1983). *The reflective practitioner : How professionals think in action*. New York: Basic Books.

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

Seeley, T. (2010). *Honeybee democracy*. Princeton N.J.: Princeton University Press.

Shaw, D., Eden, C., Ackermann, F., & School, S. G. B. (2002). Evaluating group support systems: Improving brainstorming research methodology. *RESEARCH PAPER SERIES-ASTON BUSINESS SCHOOL RESEARCH INSTITUTE*.

Sheahan, M. (2013, February 12). ThyssenKrupp warns of bleak 2013 after profit slump. Retrieved 23 April 2014, from http://www.reuters.com/article/2013/02/12/us-thyssenkrupp-results-idUSBRE91B15H20130212

Shen, H. Z., Hu, D. P., Wang, H. C., Liu, D., & Zhao, J. D. (2001). A group decision support system of distributed macroeconomics decision making. In X. F. Ju (Ed.), *PROCEEDINGS OF THE 2001 INTERNATIONAL CONFERENCE ON MANAGEMENT SCIENCE AND ENGINEERING, VOLS I AND II* (pp. 18–23). 16 FUXINGJIE NANGANGQU, HARBIN 150006, HEILONGJIANG, PEOPLES R CHINA: HARBIN INSTITUTE TECHNOLOGY PUBLISHERS.

Shleifer, A. (2000). *Inefficient markets : An introduction to behavioral finance*. Oxford; New York: Oxford University Press.

Silver, N. (2012). *The signal and the noise: why so many predictions fail--but some don't*. New York: Penguin Press.

Simon, F. B. (2013). *Gemeinsam sind wir blöd!? Die Intelligenz von Unternehmen, Managern und Märkten [Together we are stupid!? The intelligence of enterprises, managers and markets]* (4.). Heidelberg: Carl Auer.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. https://doi.org/10.1037/h0042769

Sims, R. R. (2002). Managing organizational behavior. Retrieved from http://www.netlibrary.com/urlapi.asp?action=summary&v=1&bookid=75522

Skinner, C. (2014). *Digital bank strategies to launch or become a digital bank*. Singapore: Marshall Cavendish Business. Retrieved from http://www.books24x7.com/marc.asp?bookid=65731

Slovic, S., & Slovic, P. (Eds.). (2015). *Numbers and nerves: information, emotion, and meaning in a world of data*. Corvallis, OR: Oregon State University Press.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London: W. Strahan and T. Cadell.

Sol, H. G., Takkenberg, C. A. T., & De Vries Robbé, P. F. (1987). Expert Systems and Artificial Intelligence in Decision Support Systems. In *Proceedings of the Second Mini Euroconference*. Dordrecht; Boston; Lancaster [etc.]: D. Reidel publ. Cie.

Stanzel, M. (2007). *Qualität des Aktienresearch von Finanzanalysten : Eine theoretische und empirische Untersuchung der Gewinnprognosen und Aktienempfehlungen am deutschen Kapitalmarkt [Quality of equity research of financial analysts: A theoretical and empirical investigation of earnings forecasts and stock recommendations on the German capital market]* (1. Ed.). Wiesbaden: Dt. Univ.-Verl.

Stevens, C. A., & Finlay, P. N. (1996). A research framework for group support systems. *Group Decision and Negotiation*, *5*(4), 521–543.

Stoner, J. A. F. (1961). *A comparison of individual and group decisions involving risk*. Retrieved from /z-wcorg/.

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

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

Susman, G. I. (1983). Action Research: A Sociotechnical Systems Perspective. In G. Morgan (Ed.), *Beyond method: strategies for social research*. London: SAGE Publications.

Swift, L., & Piff, S. (2005). *Quantitative methods for business, management, and finance* (2nd ed.). Houndmills Basingstoke Hampshire ;;New York: Palgrave Macmillan.

Taneja, A., & Arora, Y. (2015). The Recruitment of Intuitive Managers. *European Journal of Business and Management*, *7*(4), 305–308.

Tapscott, D., & Williams, A. (2006). *Wikinomics : How mass collaboration changes everything*. New York: Portfolio.

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

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research : Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Los Angeles: SAGE.

Tesser, A., Martin, L., & Mendolia, M. (1995). The impact of thought on attitude extremity and attitude-behavior consistency. *Conference on Attitudes and Persuasion*.

Tetlock, P. E. (2005). *Expert political judgment how good is it? How can we know?* Princeton, N.J.: Princeton University Press. Retrieved from http://public.eblib.com/EBLPublic/PublicView.do?ptiID=457930

Tetlock, P. E., & Gardner, D. (2015). *Superforecasting: the art and science of prediction* (First edition). New York: Crown Publishers.

Thaler, R. H. (2015). *Misbehaving: the making of behavioral economics* (First edition). New York: W.W. Norton & Company.

Thomas, G. (2009). *How to do your research project : A guide for students in education and applied social sciences*. London: SAGE Publications Ltd.

Thomas, G. (2011). *How to do your case study : a guide for students and researchers*. Los Angeles; London: SAGE Publications.

Thompson, K. (2016, January 15). Laboratory Experiments: Definition, Explanation, Advantages and Disadvantages. Retrieved 18 February 2017, from https://revisesociology.com/2016/01/15/laboratory-experiments-definition-explanation-advantages-and-disadvantages/

Traufetter, G. (2009). *Intuition die Weisheit der Gefühle [Intuition the Wisdom of Feelings]*. Reinbek/Hamburg: Rowohlt-Taschenbuch-Verl.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (2002). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: the psychology of intuitive judgement*. New York: Cambridge University Press.

Tziralis, G., & Tatsiopoulos, I. (2007). Prediction markets: An extended literature review. *Journal of Prediction Markets*, *1*(1), 75–91.

University of Gloucestershire. (n.d.). Doctorate in Business Administration (DBA). Retrieved 21 August 2011, from http://www.glos.ac.uk/courses/research/dba/Pages/default.aspx

University of Southampton. (n.d.). Doctorate in Business Administration (part-time). Retrieved 21 August 2011, from http://www.southampton.ac.uk/DBA

van Aken, J. E. (2001). Mode 2 knowledge production in the field of management. *Eindhoven Centre for Innovation Studies, Working Paper*, *1*.

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.

Varho, V., Rikkonen, P., & Rasi, S. (2016). Futures of distributed small-scale renewable energy in Finland — A Delphi study of the opportunities and obstacles up to 2025. *Technological Forecasting and Social Change*, *104*, 30–37. https://doi.org/10.1016/j.techfore.2015.12.001

Vincent, A., & Armstrong, M. (2010). Predicting break-points in trading strategies with Twitter.

Wallach, M., Kogan, M., & Bem, D. (1962). Group influence on individual risk taking. *Journal of Abnormal and Social Psychology*, *65*, 75–86.

Ward, A. J. W., Herbert-Read, J. E., Sumpter, D. J. T., & Krause, J. (2011). Fast and accurate decisions through collective vigilance in fish shoals. *Proceedings of the National Academy of Sciences*, *108*(6), 2312–2315. https://doi.org/10.1073/pnas.1007102108

Williams, J. B. (1938). *The theory of investment value*. Cambridge: Harvard University Press.

Yoon, K. P., & Hwang, C.-L. (1995). *Multiple attribute decision making : An introduction*. Thousand Oaks: SAGE.

Zweig, J. (2007). *Your money and your Brain : How the new science of neuroeconomics can help make you rich*. New York: Simon & Schuster.

Annex I

Collective vigilance: proportion of fish avoiding the predator

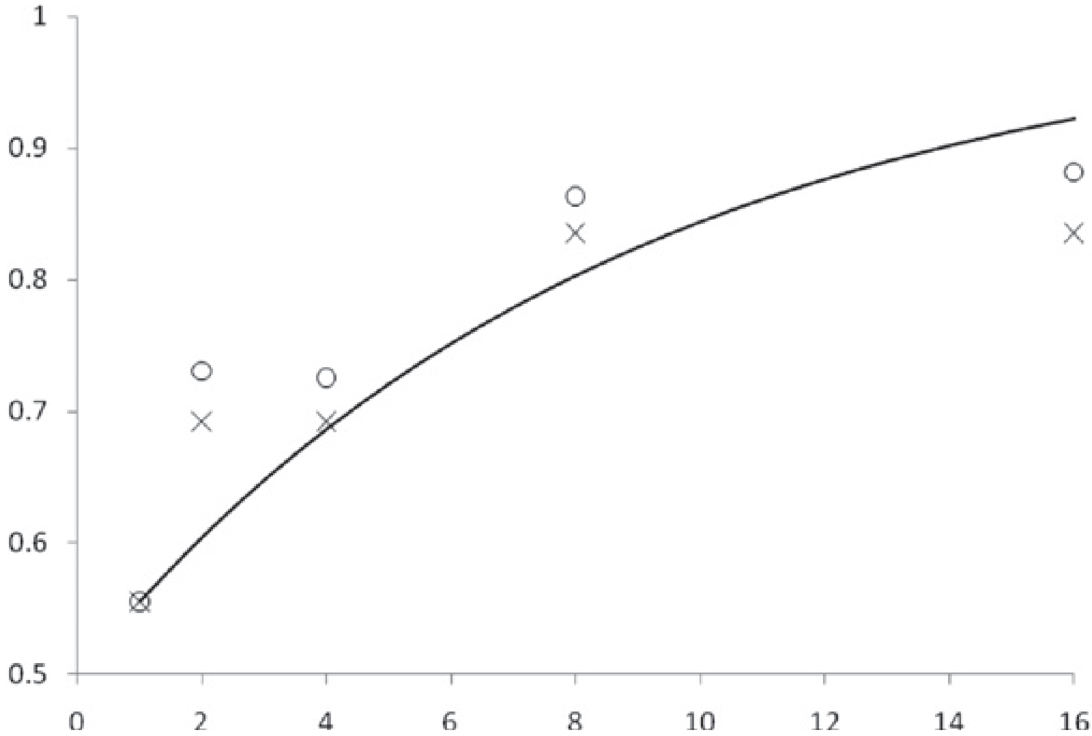


Figure 29. Collective vigilance: proportion of fish avoiding the predator

Note. Adapted from “Fast and accurate decisions through collective vigilance in fish shoals,” A. J.Ward, J. E. Herbert-Read, D. J. Sumpter, and J. Krause, *Proceedings of the National Academy of Sciences*, vol. 108, no. 6, p. 2313.

Qualitative Differences Between Five Decision Processes

Table 105. *Comparison of Qualitative Differences Between Five Decision Processes Based upon Evaluations of Leaders, Group Participants and Own Estimates*

Dimension	Interacting Groups	Nominal Groups	Delphi Technique	Prediction Markets	Collective Intelligence
Overall methodology	Unstructured face-to-face group meeting High flexibility High variability in behaviour of groups	Structured face-to-face group meeting Low flexibility Low variability in behaviour of groups	Structured series of questionnaires and feedback reports Low variability respondent behaviour	Exchange structured interactions	Unstructured interactions
Role orientation of groups	Socio-emotional Group maintenance focus	Balanced focus on social maintenance and task role	Task-instrumental focus	Task-instrumental focus	Socio-emotional Group maintenance focus
Relative quantity of ideas	Low; focused "rut" effect	Higher; independent writing and hitch-hiking round-robin	High; isolated writing of ideas	Ideas limited to pre defined trading options	Higher; independent participation and interactions
Search behaviour	Reactive search Short problem focus Task-avoidance tendency New social knowledge	Proactive search Extended problem focus High task centeredness New social and task knowledge	Proactive search Controlled problem focus High task centeredness New task knowledge	Proactive search Controlled problem focus High task centerdness	Reactive search Short problem focus Task-avoidance tendency
Normative behaviour	Conformity pressures inherent in face-to-face discussions	Tolerance for nonconformity through independent search and choice activity	Freedom not to conform through isolated anonymity	Freedom not to conform through isolated anonymity	Possible conformity pressures through group interactions
Equality of participation	Member dominance in search, evaluation, and choice phases	Member equality in search and choice phases	Respondent equality in pooling of independent judgements	Amount of participation defined by the position bought	Participation defined by activity of each group member
Method of problem solving	Person-centered Smoothing over and withdrawal	Problem-centered Conformation and problem solving	Problem-centered Majority rule of pooled independent judgements	Problem-centered Pre structured derivative	Simple group interactions
Resources utilized	Low administrative time, and cost High participant time and cost	Medium administrative time, cost, preparation High participant time and cost	High administrative	Low administrative time, and cost Medium participants time and cost	Low administrative time, and cost Medium participants time and cost
Time to obtain ideas	1 - 1/2 hours	1 - 1/2 hours	5 calendar months (shorter with e-Delphi method)	1 – 5 days	2 – 5 days

Note. Adapted from "The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes," A. H. Van de Ven & A. L. Delbecq, A. L., 1974, *Academy of Management Journal*, *17*(4), p. 618.

Theoretical Analysis of Groupthink

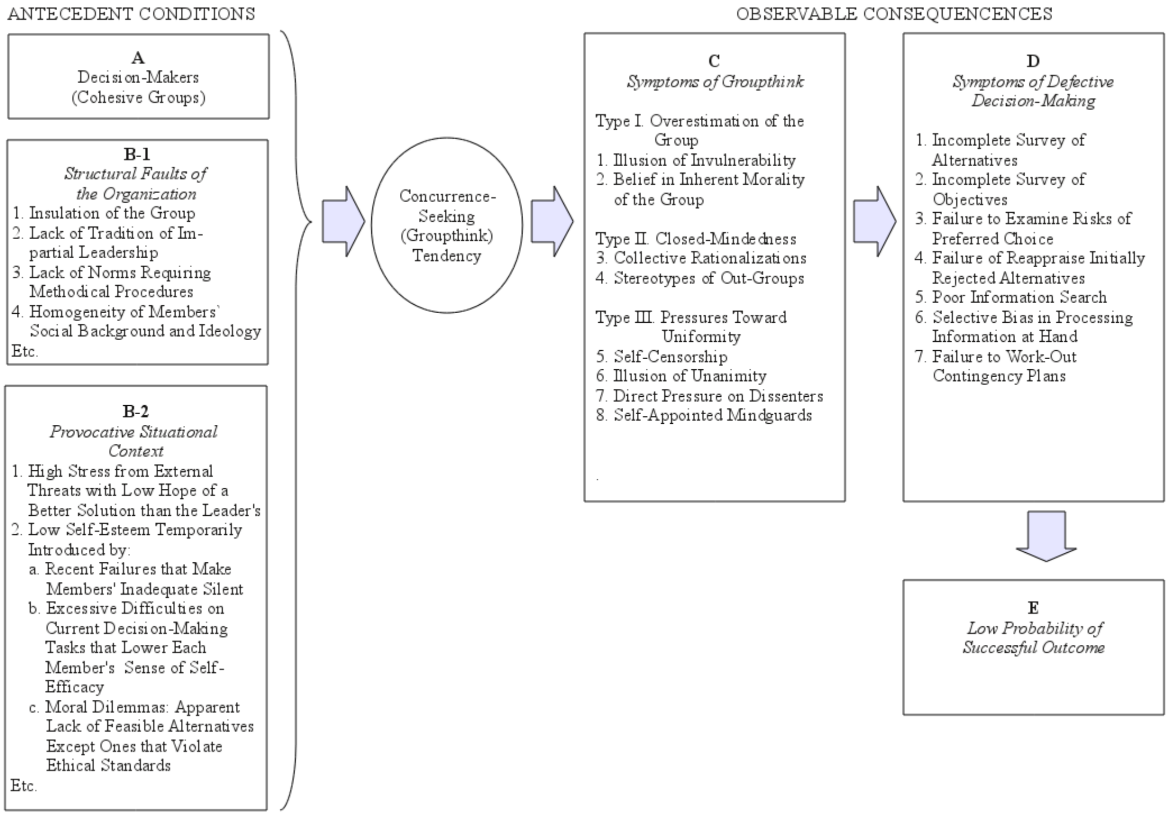


Figure 30. Theoretical Analysis of Groupthink

Note. Theoretical Analysis of Groupthink. Adapted from “Groupthink : psychological studies of policy decisions and fiascoes,” by J. L. Janis, 1982, p. 244.

Table 106. *The Continuum and Implications of Positionality in Action Research*

Insider (1) _______ (2) _______ (3) _______ (4) _______ (5) _______ (6) Outsider			
Positionality of Researcher	*Validity Criteria*	*Contributes to:*	*Traditions*
1. Insider[a] (researcher studies own self/practice)	Anderson & Herr (1999), Bullough & Pinnegar (2001), Connelly & Clandinin (1990)	Knowledge base, Improved/critiqued practice, Self/ professional transformation	Practitioner research, Autobiography, Narrative research, Self-study
2. Insider in collaboration with other insiders	Heron (1996), Saavedra (1996)	Knowledge base, Improved/critiqued practice, Professional/ organizational transformation	Feminist consciousness raising groups, Inquiry/ Study groups, Teams
3. Insider(s) in collaboration with outsider(s)	Anderson & Herr (1999), Heron (1996), Saavedra (1996)	Knowledge base, Improved/critiqued practice, Professional/ organizational transformation	Inquiry/ Study groups
4. Reciprocal collaboration (insider-outsider teams)	Anderson & Herr (1999), Bartunek & Louis (1996)	Knowledge base, Improved/critiqued practice, Professional/ organizational transformation	Collaborative forms of participatory action research that achieve equitable power relations
5. Outsider(s) in collaboration with insider(s)	Anderson & Herr (1999), Bradbury & Reason (2001), Heron (1996)	Knowledge base, Improved/critiqued practice, Organizational development/ transformation	Mainstream change agency: consultancies, industrial democracy, organizational learning; Radical change: community empowerment (Paulo Freire)
6. Outsider(s) studies insider(s)	Campbell & Stanley (1963), Lincoln & Guba (1985)	Knowledge base	University-based, academic research on action research methods or action research projects

Note. a. A flawed and deceptive version of this is when an insider studies his or her own site but fails to position himself or herself as insider to the setting *(outsider within).*
Adapted from "The action research dissertation : a guide for students and faculty," G. L. Anderson, K. Herr, 2005, p. 31.

Annex II

Data Generated with the Pilot Experiment

Comparison of Group Average Price recommendation, single expert within his narrow field of expertise (company coverage) and expert group recommendations (covered stocks and non-covered stocks combined).

Table 107. *Pilot Comparison of Group Average Price Recommendation*

Round	Date	Share_ID	Share_Name	Group Average Price	Single Expert Average Price	Expert Group Average Price	Closing Price
1	2012-02-10	1	Adidas	63.167 €	58.000 €	53.333 €	56.510 €
1	2012-02-10	2	BASF	64.256 €	55.000 €	55.333 €	60.380 €
1	2012-02-10	3	RWE	32.333 €	26.000 €	31.667 €	31.635 €
1	2012-02-10	4	ThyssenKrupp	23.256 €	26.000 €	25.667 €	21.970 €
2	2012-02-13	1	Adidas	62.625 €	65.000 €	57.333 €	57.220 €
2	2012-02-13	2	BASF	62.313 €	55.000 €	61.333 €	60.730 €
2	2012-02-13	3	RWE	37.025 €	26.000 €	31.667 €	32.160 €
2	2012-02-13	4	ThyssenKrupp	22.988 €	26.000 €	26.000 €	21.895 €
3	2012-02-17	1	Adidas	62.534 €	57.000 €	55.333 €	59.850 €
3	2012-02-17	2	BASF	63.261 €	57.000 €	55.667 €	62.980 €
3	2012-02-17	3	RWE	33.080 €	26.000 €	30.667 €	33.330 €
3	2012-02-17	4	ThyssenKrupp	22.330 €	26.000 €	24.667 €	20.550 €
4	2012-02-20	1	Adidas	63.267 €	57.000 €	56.667 €	60.340 €
4	2012-02-20	2	BASF	63.778 €	56.000 €	60.667 €	64.660 €
4	2012-02-20	3	RWE	35.056 €	26.000 €	32.333 €	33.660 €
4	2012-02-20	4	ThyssenKrupp	21.789 €	25.000 €	24.667 €	21.370 €
5	2012-02-24	1	Adidas	62.200 €	62.000 €	58.000 €	58.750 €
5	2012-02-24	2	BASF	64.700 €	68.000 €	68.500 €	64.700 €
5	2012-02-24	4	ThyssenKrupp	21.633 €	26.000 €	24.500 €	20.535 €
6	2012-02-27	1	Adidas	61.300 €	65.000 €	56.667 €	58.200 €
6	2012-02-27	2	BASF	66.229 €	70.000 €	66.000 €	65.810 €
6	2012-02-27	3	RWE	33.586 €	28.000 €	33.000 €	33.465 €
6	2012-02-27	4	ThyssenKrupp	22.064 €	25.000 €	25.000 €	20.355 €
7	2012-03-02	1	Adidas	62.638 €	65.000 €	58.333 €	59.330 €
7	2012-03-02	2	BASF	66.100 €	68.000 €	65.333 €	66.640 €
7	2012-03-02	3	RWE	38.675 €	38.000 €	37.000 €	34.730 €
7	2012-03-02	4	ThyssenKrupp	21.175 €	25.000 €	25.000 €	20.380 €
8	2012-03-05	1	Adidas	61.543 €	65.000 €	58.333 €	59.470 €
8	2012-03-05	2	BASF	66.329 €	69.000 €	65.667 €	65.650 €
8	2012-03-05	3	RWE	35.317 €	38.000 €	37.333 €	34.495 €
8	2012-03-05	4	ThyssenKrupp	22.538 €	25.000 €	25.333 €	20.000 €
9	2012-03-09	1	Adidas	61.370 €	62.000 €	57.333 €	58.100 €
9	2012-03-09	2	BASF	62.990 €	69.000 €	66.667 €	65.220 €
9	2012-03-09	3	RWE	34.110 €	30.000 €	34.667 €	35.270 €
9	2012-03-09	4	ThyssenKrupp	20.800 €	23.000 €	24.000 €	19.350 €

Table 108. Group of laypeople

Round	Share ID	Share Name	N	Sum Price	Average Price	Std Price	Max Rec	Min Rec	Range	Variance Price
1	1	Adidas	9	568.500 €	63.167 €	8.172 €	85.000 €	55.000 €	30.000 €	66.778 €
1	2	BASF	9	578.300 €	64.256 €	11.097 €	90.000 €	50.000 €	40.000 €	123.145 €
1	3	RWE	9	291.000 €	32.333 €	5.292 €	42.000 €	25.000 €	17.000 €	28.000 €
1	4	ThyssenKrupp	9	209.300 €	23.256 €	3.466 €	27.500 €	17.000 €	10.500 €	12.011 €
2	1	Adidas	8	501.000 €	62.625 €	1.867 €	65.000 €	60.000 €	5.000 €	3.484 €
2	2	BASF	8	498.500 €	62.313 €	5.994 €	70.000 €	50.000 €	20.000 €	35.934 €
2	3	RWE	8	296.200 €	37.025 €	9.354 €	60.000 €	28.000 €	32.000 €	87.492 €
2	4	ThyssenKrupp	8	183.900 €	22.988 €	3.678 €	28.000 €	17.000 €	11.000 €	13.526 €
3	1	Adidas	10	625.340 €	62.534 €	3.177 €	67.000 €	55.000 €	12.000 €	10.094 €
3	2	BASF	10	632.610 €	63.261 €	4.373 €	70.000 €	55.000 €	15.000 €	19.124 €
3	3	RWE	10	330.800 €	33.080 €	3.790 €	40.500 €	28.000 €	12.500 €	14.366 €
3	4	ThyssenKrupp	10	223.300 €	22.330 €	3.313 €	29.500 €	18.400 €	11.100 €	10.978 €
4	1	Adidas	9	569.400 €	63.267 €	2.572 €	66.000 €	57.000 €	9.000 €	6.613 €
4	2	BASF	9	574.000 €	63.778 €	3.945 €	70.000 €	58.000 €	12.000 €	15.562 €
4	3	RWE	9	315.500 €	35.056 €	3.557 €	40.500 €	30.900 €	9.600 €	12.649 €
4	4	ThyssenKrupp	9	196.100 €	21.789 €	3.840 €	29.500 €	18.000 €	11.500 €	14.745 €
5	1	Adidas	9	559.800 €	62.200 €	3.688 €	67.000 €	55.000 €	12.000 €	13.598 €
5	2	BASF	9	582.300 €	64.700 €	5.265 €	72.500 €	55.000 €	17.500 €	27.720 €
5	3	RWE	9	290.900 €	32.322 €	2.904 €	38.000 €	28.000 €	10.000 €	8.435 €
5	4	ThyssenKrupp	9	194.700 €	21.633 €	3.552 €	27.500 €	18.000 €	9.500 €	12.620 €
6	1	Adidas	7	429.100 €	61.300 €	3.373 €	65.000 €	55.000 €	10.000 €	11.374 €
6	2	BASF	7	463.600 €	66.229 €	3.834 €	72.500 €	62.000 €	10.500 €	14.699 €
6	3	RWE	7	235.100 €	33.586 €	2.874 €	39.500 €	30.000 €	9.500 €	8.261 €
6	4	ThyssenKrupp	7	154.450 €	22.064 €	3.780 €	28.500 €	19.000 €	9.500 €	14.291 €
7	1	Adidas	8	501.100 €	62.638 €	1.608 €	65.000 €	60.800 €	4.200 €	2.585 €
7	2	BASF	7	462.700 €	66.100 €	2.918 €	70.000 €	62.000 €	8.000 €	8.517 €
7	3	RWE	8	309.400 €	38.675 €	12.119 €	69.100 €	28.000 €	41.100 €	146.882 €
7	4	ThyssenKrupp	8	169.400 €	21.175 €	3.403 €	28.500 €	18.000 €	10.500 €	11.582 €
8	1	Adidas	7	430.800 €	61.543 €	2.038 €	64.300 €	58.000 €	6.300 €	4.154 €
8	2	BASF	7	464.300 €	66.329 €	2.508 €	69.500 €	62.000 €	7.500 €	6.291 €
8	3	RWE	6	211.900 €	35.317 €	2.672 €	40.500 €	32.600 €	7.900 €	7.141 €
8	4	ThyssenKrupp	8	180.300 €	22.538 €	5.493 €	34.500 €	18.000 €	16.500 €	30.170 €
9	1	Adidas	10	613.700 €	61.370 €	3.054 €	67.400 €	56.000 €	11.400 €	9.328 €
9	2	BASF	10	629.900 €	62.990 €	9.502 €	69.500 €	35.000 €	34.500 €	90.295 €
9	3	RWE	10	341.100 €	34.110 €	4.009 €	41.000 €	28.000 €	13.000 €	16.073 €
9	4	ThyssenKrupp	10	208.000 €	20.800 €	3.487 €	27.500 €	17.000 €	10.500 €	12.158 €
10	1	Adidas	9	534.900 €	59.433 €	4.039 €	65.000 €	50.000 €	15.000 €	16.313 €
10	2	BASF	9	602.200 €	66.911 €	4.374 €	75.000 €	58.000 €	17.000 €	19.128 €
10	3	RWE	9	321.200 €	35.689 €	3.549 €	41.000 €	29.000 €	12.000 €	12.597 €
10	4	ThyssenKrupp	9	179.000 €	19.889 €	3.644 €	27.500 €	15.000 €	12.500 €	13.279 €

Table 109. *(Small) group of experts*

Round	Share ID	Share Name	N	Sum Price	Average Price	Std Price	Max Rec	Min Rec	Range	Variance Price
1	1	Adidas	3	160.000 €	53.333 €	3.399 €	58.000 €	50.000 €	8.000 €	11.556 €
1	2	BASF	3	166.000 €	55.333 €	0.471 €	56.000 €	55.000 €	1.000 €	0.222 €
1	3	RWE	3	95.000 €	31.667 €	4.028 €	35.000 €	26.000 €	9.000 €	16.222 €
1	4	ThyssenKrupp	3	77.000 €	25.667 €	0.471 €	26.000 €	25.000 €	1.000 €	0.222 €
2	1	Adidas	3	172.000 €	57.333 €	5.558 €	65.000 €	52.000 €	13.000 €	30.889 €
2	2	BASF	3	184.000 €	61.333 €	4.497 €	65.000 €	55.000 €	10.000 €	20.222 €
2	3	RWE	3	95.000 €	31.667 €	4.028 €	35.000 €	26.000 €	9.000 €	16.222 €
2	4	ThyssenKrupp	3	78.000 €	26.000 €	0.000 €	26.000 €	26.000 €	0.000 €	0.000 €
3	1	Adidas	3	166.000 €	55.333 €	1.247 €	57.000 €	54.000 €	3.000 €	1.556 €
3	2	BASF	3	167.000 €	55.667 €	0.943 €	57.000 €	55.000 €	2.000 €	0.889 €
3	3	RWE	3	92.000 €	30.667 €	3.682 €	35.000 €	26.000 €	9.000 €	13.556 €
3	4	ThyssenKrupp	3	74.000 €	24.667 €	0.943 €	26.000 €	24.000 €	2.000 €	0.889 €
4	1	Adidas	3	170.000 €	56.667 €	1.247 €	58.000 €	55.000 €	3.000 €	1.556 €
4	2	BASF	3	182.000 €	60.667 €	5.249 €	68.000 €	56.000 €	12.000 €	27.556 €
4	3	RWE	3	97.000 €	32.333 €	4.643 €	37.000 €	26.000 €	11.000 €	21.556 €
4	4	ThyssenKrupp	3	74.000 €	24.667 €	0.471 €	25.000 €	24.000 €	1.000 €	0.222 €
5	1	Adidas	2	116.000 €	58.000 €	4.000 €	62.000 €	54.000 €	8.000 €	16.000 €
5	2	BASF	2	137.000 €	68.500 €	0.500 €	69.000 €	68.000 €	1.000 €	0.250 €
5	3	RWE	2	71.000 €	35.500 €	0.500 €	36.000 €	35.000 €	1.000 €	0.250 €
5	4	ThyssenKrupp	2	49.000 €	24.500 €	1.500 €	26.000 €	23.000 €	3.000 €	2.250 €
6	1	Adidas	3	170.000 €	56.667 €	6.236 €	65.000 €	50.000 €	15.000 €	38.889 €
6	2	BASF	3	198.000 €	66.000 €	4.320 €	70.000 €	60.000 €	10.000 €	18.667 €
6	3	RWE	3	99.000 €	33.000 €	3.559 €	36.000 €	28.000 €	8.000 €	12.667 €
6	4	ThyssenKrupp	3	75.000 €	25.000 €	0.816 €	26.000 €	24.000 €	2.000 €	0.667 €
7	1	Adidas	3	175.000 €	58.333 €	4.714 €	65.000 €	55.000 €	10.000 €	22.222 €
7	2	BASF	3	196.000 €	65.333 €	3.771 €	68.000 €	60.000 €	8.000 €	14.222 €
7	3	RWE	3	111.000 €	37.000 €	1.414 €	38.000 €	35.000 €	3.000 €	2.000 €
7	4	ThyssenKrupp	3	75.000 €	25.000 €	0.816 €	26.000 €	24.000 €	2.000 €	0.667 €
8	1	Adidas	3	175.000 €	58.333 €	4.714 €	65.000 €	55.000 €	10.000 €	22.222 €
8	2	BASF	3	197.000 €	65.667 €	4.028 €	69.000 €	60.000 €	9.000 €	16.222 €
8	3	RWE	3	112.000 €	37.333 €	0.943 €	38.000 €	36.000 €	2.000 €	0.889 €
8	4	ThyssenKrupp	3	76.000 €	25.333 €	0.471 €	26.000 €	25.000 €	1.000 €	0.222 €
9	1	Adidas	3	172.000 €	57.333 €	3.300 €	62.000 €	55.000 €	7.000 €	10.889 €
9	2	BASF	3	200.000 €	66.667 €	3.300 €	69.000 €	62.000 €	7.000 €	10.889 €
9	3	RWE	3	104.000 €	34.667 €	3.399 €	38.000 €	30.000 €	8.000 €	11.556 €
9	4	ThyssenKrupp	3	72.000 €	24.000 €	2.160 €	27.000 €	22.000 €	5.000 €	4.667 €
10	1	Adidas	3	169.000 €	56.333 €	4.190 €	62.000 €	52.000 €	10.000 €	17.556 €
10	2	BASF	3	197.000 €	65.667 €	4.028 €	69.000 €	60.000 €	9.000 €	16.222 €
10	3	RWE	3	94.000 €	31.333 €	3.399 €	36.000 €	28.000 €	8.000 €	11.556 €
10	4	ThyssenKrupp	3	75.000 €	25.000 €	2.449 €	28.000 €	22.000 €	6.000 €	6.000 €

Table 110. *Single expert within his narrow field of expertise (recommendation for covered shares)*

Round	Share ID	Share Name	N	Sum Price	Average Price	Std Price	Max Rec	Min Rec
1	1	Adidas	1	€58.000	€58.000	€0.000	€58.000	€58.000
1	2	BASF	1	€55.000	€55.000	€0.000	€55.000	€55.000
1	3	RWE	1	€26.000	€26.000	€0.000	€26.000	€26.000
1	4	ThyssenKrupp	1	€26.000	€26.000	€0.000	€26.000	€26.000
2	1	Adidas	1	€65.000	€65.000	€0.000	€65.000	€65.000
2	2	BASF	1	€55.000	€55.000	€0.000	€55.000	€55.000
2	3	RWE	1	€26.000	€26.000	€0.000	€26.000	€26.000
2	4	ThyssenKrupp	1	€26.000	€26.000	€0.000	€26.000	€26.000
3	1	Adidas	1	€57.000	€57.000	€0.000	€57.000	€57.000
3	2	BASF	1	€57.000	€57.000	€0.000	€57.000	€57.000
3	3	RWE	1	€26.000	€26.000	€0.000	€26.000	€26.000
3	4	ThyssenKrupp	1	€26.000	€26.000	€0.000	€26.000	€26.000
4	1	Adidas	1	€57.000	€57.000	€0.000	€57.000	€57.000
4	2	BASF	1	€56.000	€56.000	€0.000	€56.000	€56.000
4	3	RWE	1	€26.000	€26.000	€0.000	€26.000	€26.000
4	4	ThyssenKrupp	1	€25.000	€25.000	€0.000	€25.000	€25.000
5	1	Adidas	1	€62.000	€62.000	€0.000	€62.000	€62.000
5	2	BASF	1	€68.000	€68.000	€0.000	€68.000	€68.000
5	4	ThyssenKrupp	1	€26.000	€26.000	€0.000	€26.000	€26.000
6	1	Adidas	1	€65.000	€65.000	€0.000	€65.000	€65.000
6	2	BASF	1	€70.000	€70.000	€0.000	€70.000	€70.000
6	3	RWE	1	€28.000	€28.000	€0.000	€28.000	€28.000
6	4	ThyssenKrupp	1	€25.000	€25.000	€0.000	€25.000	€25.000
7	1	Adidas	1	€65.000	€65.000	€0.000	€65.000	€65.000
7	2	BASF	1	€68.000	€68.000	€0.000	€68.000	€68.000
7	3	RWE	1	€38.000	€38.000	€0.000	€38.000	€38.000
7	4	ThyssenKrupp	1	€25.000	€25.000	€0.000	€25.000	€25.000
8	1	Adidas	1	€65.000	€65.000	€0.000	€65.000	€65.000
8	2	BASF	1	€69.000	€69.000	€0.000	€69.000	€69.000
8	3	RWE	1	€38.000	€38.000	€0.000	€38.000	€38.000
8	4	ThyssenKrupp	1	€25.000	€25.000	€0.000	€25.000	€25.000
9	1	Adidas	1	€62.000	€62.000	€0.000	€62.000	€62.000
9	2	BASF	1	€69.000	€69.000	€0.000	€69.000	€69.000
9	3	RWE	1	€30.000	€30.000	€0.000	€30.000	€30.000
9	4	ThyssenKrupp	1	€23.000	€23.000	€0.000	€23.000	€23.000
10	1	Adidas	1	€62.000	€62.000	€0.000	€62.000	€62.000
10	2	BASF	1	€69.000	€69.000	€0.000	€69.000	€69.000
10	3	RWE	1	€30.000	€30.000	€0.000	€30.000	€30.000
10	4	ThyssenKrupp	1	€25.000	€25.000	€0.000	€25.000	€25.000

One Week Predictions Pilot Stage

The analysis of the short-term predictions (1 week) were analysed and preliminary results and findings were presented at the 2nd Annual Doctoral Colloquium in Berlin on July 14 (Endress, 2012); the results for 3-month predictions could not be presented there because at the hand-in date for colloquium papers the 3-month period was not over and, accordingly, it was still unclear whether the predictions would turn out to be right or wrong. However, the results of the 3 month period, as well as a summary of the one week results, were presented at the 10th International CIRCLE conference (Endress & Gear, 2013b). Some of the preliminary results of the analysis of the pilot experiment and the secondary data gathered from existing communities have also been published in peer reviewed journals (Endress, 2013; Endress & Gear, 2013a).

The examination of the first estimates (for one week) showed that the group of lay people was slightly better at predicting stock price movements than the experts (see Table 111). 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), and these undecided rounds were excluded from the analysis. The group’s outperformance was even higher in weeks when the stock price was declining. From 17 predictions (m=17), the group had 10 (71.4%) correct predictions (three undecided rounds were excluded), the expert group had six (35.3%) correct predictions and the single expert had nine (52.9%) correct predictions.

Table 111. *Aggregated 1 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 112), even though almost all group members stated in the interviews that they were not influenced by the group feedback from the e-Delphi rounds. That the group had a tendency towards conforming might be possible, in particular with price predictions. Additional data might help to gain more knowledge about that process.

Table 112. *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

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

Table 113. *Pilot Run One Week Predictions of Lay Participants*

	right	wrong	Measurements	Success Rate	Skill (Self Est.)
Participant 1	17	15	32	53.10%	3
Participant 2	20	16	36	55.60%	3-4
Participant 3	20	12	32	62.50%	2-3
Participant 4	22	14	36	61.10%	6
Participant 5	23	13	36	63.90%	1
Participant 6	22	18	40	55.00%	2
Participant 7	22	18	40	55.00%	7
Participant 8	26	10	36	72.20%	2
Participant 9	17	14	31	54.80%	7-8
Participant 10	17	15	32	53.10%	5
Participant 11	14	18	32	43.80%	2
Ø				57.30%	

At this stage it was only possible to examine the one-week predictions. Further analysis, such as comparison of the three-month estimates, was not possible until the actual stock price at the end of this period was available. All participants in 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 able to give an estimate or at least enter a guess as to whether the stock price was going up or down. One participant felt uncomfortable about 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, 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 not as mandatory fields in the online survey, but rather to leave it to the participants to enter a concrete price target.

Three Month Predictions Pilot Stage

The 3-month predictions consist basically of two components: The first component is an estimate of whether the share would go up or down, and the second component is an actual target price estimate for a 3-month period. Every participant had to enter both these components independently for the four stocks in the pilot experiment.

Accuracy of Individual Predictions of e-Delphi Group Members

Table 114. *Results Overview: 3-Month Predictions of Lay Participants*

3 Months	Right	Wrong	Measurements	Success Rate	Skill (Self-Est.)
Participant 1	20	12	32	62.5%	3
Participant 2	23	13	36	63.9%	3–4
Participant 3	22	10	32	68.8%	2–3
Participant 4	10	26	36	27.8%	6
Participant 5	16	20	36	44.4%	1
Participant 6	20	20	40	50.0%	2
Participant 7	26	6	32	81.3%	7
Participant 8	22	14	36	61.1%	2
Participant 9	20	12	32	62.5%	7–8
Participant 10	12	20	32	37.5%	5
Participant 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 114). 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 were interviewed in parallel to the e-Delphi rounds. In the interviews, all participants were asked to give a self-assessment of their investment expertise on a scale from 1 to 10 (1 = no knowledge; 10 = expert). It might be hypothesized that there would be 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 assumption, for 3-month predictions, there was a correlation of 0.12 and even a slightly negative correlation of self-estimated skill and success rate for the 1-week predictions (-0.20). Table 116 shows the predictive accuracy of the individual predictions of the experts (professional financial analysts) for the 3-month estimates.

Comparison of 3-month predictions per share.

In the examination of the longer-term estimates (for 3 months), the group of lay people was again better at predicting the stock price movement than were the experts (see Table 115). Of 40 predictions (m = 40), the group had 17 predictions 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 115. *Comparison of 3-Month Predictions per Share*

3-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 3-month predictions per share (see Table 115) shows that the lay group had more correct predictions than the expert group and also slightly more than the single expert within his narrow field of expertise. This result contradicts the assumption that while lay people might guess the price movement and sentiment more correctly in the short term, on a period longer than a week, 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 116).

Table 116. *Results Overview: 3-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 3-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 is convergence of answers and an improvement in the 2nd 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 (Fischer, 1978; Linstone & Turoff, 2002; van de Ven & Delbecq, 1974). Despite all the controversy about the correct application and 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 towards conforming with the group opinion in the second Delphi round (Fischer, 1978; Linstone & Turoff, 2002; Rowe & Wright, 1999; Rowe et al., 2005).

Table 117. *Comparison of 3-Month Group 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

One should note that many studies using the Delphi method have no stringent follow-up, and it is often unclear whether the predictions made using 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 117) had a follow-up, and even though it was only three months later, it is possible to assess whether the predictions were correct or not. The pilot experiment contradicts the view that there is an improvement with the second Delphi round. This might be attributed to the research design and feedback loop. Since participants get information about share prices and company development not only from the Delphi group but also from other sources, it might be possible that they rely more on the information from outside the group. In interviews with the group participants, it was also mentioned that they did not read the feedback provided before giving the second prediction. Some participants also mentioned that they did not trust the group because they did not know the expertise of the group participants or their rationale for the prediction. Linstone and Turoff already 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.

Table 118. *Comparison of 1-Week Predictions from Week 1-5 and Week 6-10*

1 Week	Single Expert		Expert Group		Lay Group			Measurements
e-Delphi	right	wrong	right	wrong	right	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 119. *Comparison of 3-Month Predictions from Week 1-5 and Week 6-10*

3 Month	Single Expert		Expert Group		Lay Group			Measurements
e-Delphi	right	wrong	right	wrong	right	wrong	excluded	
Weeks 1-5	11	9	7	13	7	10	3	20
Weeks 6-10	4	16	3	17	10	9	1	20

Changes recommended by lay group participants.

An analysis of the change behaviour of the participants shows that participants 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 were 56 actual changes of prediction during the pilot (see Table 120). That means that 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 only 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 in 3-month predictions did not bring such a big improvement: Of 27 changes, 14 were correct and 13 wrong. This means that for the 3-month predictions there was an overall improvement of only one recommendation.

Table 120. *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 129 in the appendix) shows that there was not a big difference in prediction accuracy overall between the lay group and expert price estimates. The price estimate from the lay group averaged 17.58% off the target from actual market price, the single expert 17.41%, and the expert group 17.63%. On the level of individual shares, there were still some major differences: The single expert was better for RWE 3-month price estimates, and the group outperformed the experts in the case of ThyssenKrupp (see Table 129). This finding goes well with the assumption and observation that the lay group does perform well in comparison with experts, especially in the case of falling stock prices. ThyssenKrupp lost about 40% of market value (see Table 112), by far the highest loss of all shares in the pilot experiment.

Price movement changes

An analysis of the change events during the 1 week predictions showed that there were 9 changes of direction (in terms of movement change from up/down) in the 5 weeks of the pilot. Adidas changed price movement direction 4 times, Bayer changed direction 4 times, ThyssenKrupp changed twice, and RWE changed twice. These 6 changes were correctly predicted by the group of experts 6 times, by the single experts 6 times and by the lay group 5 times.

An analysis of the change events during the 3 month predictions showed that there were only 6 changes of direction in the 5 weeks. Bayer and ThyssenKrupp did not change, but were always going down. RWE changed once and Adidas changed direction 5 times. These 6 changes were correctly predicted by the group of experts 3 times, by the single experts 3 times and by the lay group 2 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.

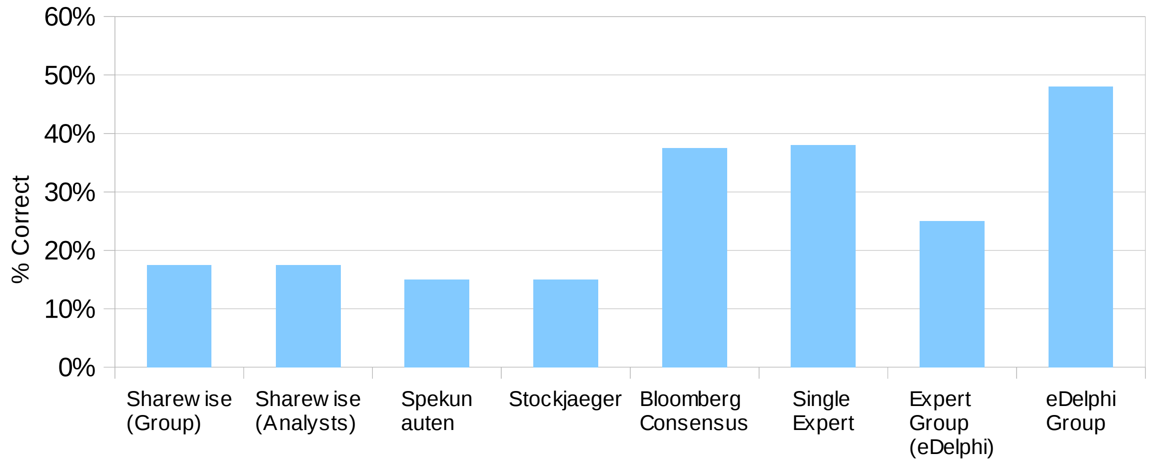


Figure 31. Group comparison: 3-month performance.

Stock Trading Communities (Pilot Run)

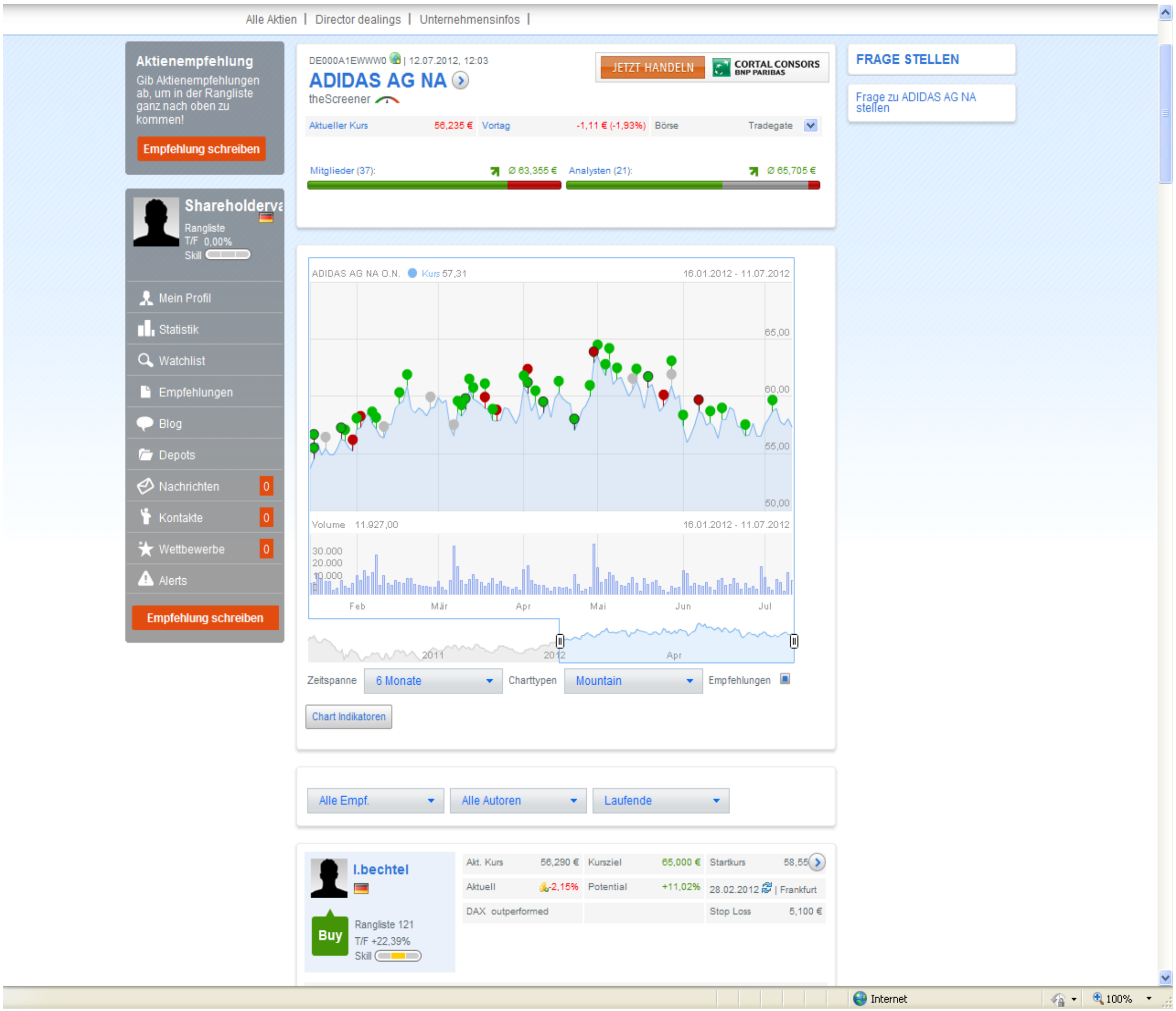


Figure 32: Example Screenshot from Sharewise Community

Table 121. *Sharewise Community Results (Pilot Run)*

http://www.de.sharewise.com/									
	11.02.12	**13.02.12**	**17.02.12**	**24.02.12**	**27.02.12**	**02.03.12**	**05.03.12**	**09.03.12**	**12.03.12**
ADIDAS AG. (DE000A1EWWW0)									
Member	29	29	29	27	27	27	26	28	29
Buy	24	24	24	22	22	22	21	23	24
Sell	5	5	5	5	5	5	5	5	5
Ø-Target Price	64.666 €	64.666 €	64.666 €	65.607 €	65.607 €	65.748 €	65.584 €	65.578 €	65.558 €
Analysts	9	9	10	9	8	9	9	14	14
Buy	5	5	6	5	4	4	4	7	7
Hold	4	4	4	4	4	5	5	7	7
Sell	0	0	0	0	0	0			
Ø-Target Price	61.711 €	61.711 €	62.040 €	62.267 €	61.863 €	64.656 €	61.656 €	64.354 €	64.354 €
BASF SE (DE000BASF111)									
Mitglieder	34	31	31	30	32	31	29	30	31
Buy	25	22	22	21	23	23	20	21	22
Sell	9	9	9	9	9	9	9	9	9
Ø-Target Price	57.684 €	57.137 €	57.137 €	57.771 €	58.497 €	59.185 €	59.787 €	65.600 €	61.059 €
Analysten	15	15	16	17	21	20	19	20	20
Buy	6	6	7	8	11	11	10	11	11
Hold	8	8	8	8	9	8	8	8	8
Sell	1	1	1	1	1	1	1	1	1
Ø-Target Price	61.643 €	61.643 €	62.800 €	61.786 €	65.474 €	65.944 €	65.706 €	66.778 €	66.778 €
RWE (DE0007037129)									
Mitglieder	20	19	17	17	17	15	16	15	15
Buy	15	14	12	12	12	10	10	9	9
Sell	5	5	5	5	5	6	6	6	6
Ø-Target Price	31.549 €	31.677 €	31.198 €	31.665 €	32.253 €	32.400 €	32.400 €	33.293 €	33.293 €
Analysten	21	21	20	21	21	19	19	22	22
Buy	6	6	6	6	6	5	5	8	8
Hold	10	10	9	10	10	9	9	9	9
Sell	5	5	5	5	5	5	5	5	5
Ø-Target Price	31.925 €	31.925 €	31.658 €	31.800 €	31.800 €	31.889 €	31.889 €	34.619 €	34.619 €
THYSSENKRUPP. (DE0007500001)									
Mitglieder	50	43	41	41	39	42	41	40	40
Buy	42	35	33	33	31	31	30	30	30
Sell	8	8	8	8	8	11	11	10	10
Ø-Target Price	22.200 €	25.528 €	24.630 €	24.459 €	24.021 €	23.305 €	23.313 €	23.458 €	23.458 €
Analysten	20	20	19	19	19	19	19	19	19
Buy	9	9	9	10	10	10	10	10	10
Hold	10	10	10	8	8	8	8	8	8
Sell	1	1	0	1	1	1	1	1	1
Ø-Target Price	24.447 €	24.447 €	24.989 €	25.150 €	25.150 €	25.150 €	25.150 €	25.150 €	25.150 €

Table 122. *Spekunauten Community Results (Pilot Run)*

http://www.spekunauten.de/

	11.02.12	**13.02.12**	**17.02.12**	**24.02.12**	**27.02.12**	**02.03.12**	**05.03.12**	**09.03.12**	**12.03.12**
ADIDAS AG. (DE000A1EWWW0)									
Members	26	26	27	28	28	28	28	28	28
Buy	26	26	27	28	28	28	28	28	28
Sell	0	0	0	0	0	0	0	0	0
BASF SE (DE000BASF111)									
Members	28	28	28	29	29	29	29	29	29
Buy	26	26	26	27	27	27	27	27	27
Sell	2	2	2	2	2	2	2	2	2
RWE (DE0007037129)									
Members	28	28	28	28	28	28	28	28	28
Buy	14	14	14	14	14	14	14	14	14
Sell	14	14	14	14	14	14	14	14	14
THYSSENKRUPP. (DE0007500001)									
Members	25	25	25	25	25	25	25	25	25
Buy	20	20	20	20	20	20	20	20	20
Sell	5	5	5	5	5	5	5	5	5

Table 123. *Stockjaeger Community Results (Pilot Run)*

http://www.stockjaeger.de/									
	11.02.12	**13.02.12**	**17.02.12**	**24.02.12**	**27.02.12**	**02.03.12**	**05.03.12**	**09.03.12**	**12.03.12**
ADIDAS AG. (DE000A1EWWW0)									
Members	2	2	2	2	2	2	2	2	2
Buy	2	2	2	2	2	2	2	2	2
Sell	0	0	0	0	0	0	0	0	0
Ø-Target Price	44.500 €	44.500 €	44.500 €	44.500 €	44.500 €	44.500 €	44.500 €	44.500 €	44.500 €
BASF SE (DE000BASF111)									
Members	0	0	0	0	0	0	0	0	0
Buy	0	0	0	0	0	0	0	0	0
Sell	0.000 €	0.000 €	0.000 €	0.000 €	0.000 €	0.000 €	0.000 €	0.000 €	0.000 €
Ø-Target Price	-	-	-	-	-	-	-	-	-
RWE (DE0007037129)									
Members	2	2	2	2	2	2	2	2	2
Buy	1	1	1	1	1	1	1	1	1
Sell	1	1	1	1	1	1	1	1	1
Ø-Target Price	45.000 €	45.000 €	45.000 €	45.000 €	45.000 €	45.000 €	45.000 €	45.000 €	45.000 €
THYSSENKRUPP. (DE0007500001)									
Members	2	2	2	2	2	2	2	2	2
Buy	2	2	2	2	2	2	2	2	2
Sell	0	0	0	0	0	0	0	0	0
Ø-Target Price	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €
	Only one target price recom-mendation available	Only one target price recom-mendation available	Only one target price recom-mendation available	Only one target price recom-mendation available	Only one target price recom-mendation available	Only one target price recom-mendation available	Only one target price recom-mendation available	Only one target price recom-mendation available	Only one target price recom-mendation available

Stock Trading Communities (Main Experiment)

Table 124. *Sharewise Community Results (Main Experiment) part 1*

	http://www.de.sharewise.com/									
	17/11/2012	**19/11/12**	**24/11/12**	**26/11/12**	**01/12/12**	**03/12/12**	**08/12/12**	**10/12/12**	**15/12/12**	**17/12/12**
ADIDAS AG. (DE000A1EWWW0)										
Member	11	11	12	11	11	11	10	10	10	10
Buy	4	4	5	5	5	5	5	5	5	5
Sell	7	7	7	6	6	6	5	5	5	5
Ø-TP	56.8	56.8	57.983	58.345	58.345	58.345	59.98	59.98	61.08	61.08
Analysts	16	16	16	16	16	16	17	17	17	17
Buy	9	9	9	9	9	9	10	10	10	10
Hold	6	6	6	6	6	6	6	6	6	6
Sell	1	1	1	1	1	1	1	1	1	1
Ø-TP	62.1	62.1	62.1	62.1	62.1	62.1	62.1	62.1	62.1	62.1
HEIDELBERGCEMENT (DE0006047004)										
Member	7	7	6	6	6	7	7	8	9	9
Buy	3	3	3	3	3	3	3	4	5	5
Sell	4	4	3	3	3	4	4	4	5	5
Ø-TP	38.286	38.286	39.667	39.667	39.667	39.571	39.571	40.275	40.929	40.929
Analysts	18	18	18	18	18	18	18	17	17	17
Buy	10	10	10	10	11	11	11	10	10	10
Hold	4	4	4	4	3	3	3	3	3	3
Sell	4	4	4	4	4	4	4	4	4	4
Ø-TP	45	45	45	45	42.75	42.75	42.75	40.667	40.667	40.667
RWE (DE0007037129)										
Member	11	13	14	14	13	13	13	13	13	13
Buy	10	11	11	11	10	10	10	10	10	10
Sell	2	2	3	3	3	3	3	3	3	3
Ø-TP	39.13	38.735	37.076	37.076	36.851	36.851	37.08	37.06	37.06	37.06
Analysts	21	21	22	22	22	22	22	22	22	22
Buy	7	7	6	6	6	6	6	6	6	6
Hold	10	10	12	12	12	12	12	11	9	9
Sell	4	4	4	4	4	4	4	5	7	7
Ø-TP	37.25	37.25	35.667	35.667	35.667	35.667	35.667	38.5	40	40
Siemens (DE0007236101)										
Member	27	27	26	25	22	22	22	23	24	24
Buy	15	15	14	13	11	11	11	11	12	12
Sell	12	12	12	12	11	11	11	12	12	12
Ø-TP	72.527	72.527	72.898	72.674	72.811	72.811	72.811	71.558	72.202	72.202
Analysts	19	19	19	19	19	19	19	19	19	19
Buy	4	4	4	4	4	4	4	4	4	4
Hold	14	14	13	13	13	13	13	13	13	13
Sell	1	1	2	2	2	2	2	2	2	2
Ø-TP	90	90	90	90						
THYSSENKRUPP. (DE0007500001)										
Member	30	30	29	30	29	30	32	30	30	32
Buy	21	21	20	19	18	18	20	21	21	16
Sell	9	9	9	11	12	12	12	9	9	16
Ø-TP	17.08	17.08	16.98	16.701	16.644	16.586	16.699	17.08	17.08	16.505
Analysts	17	17	17	17	18	18	18	17	17	19
Buy	7	7	7	6	6	5	5	7	7	6
Hold	6	6	6	7	8	8	8	6	6	8
Sell	4	4	4	4	4	5	5	4	4	5
Ø-TP	16.93	16.93	17	16.429	16.429	15.833	15.833	16.93	16.93	16.625

Table 125. *Sharewise Community Results (Main Experiment) part 2*

	http://www.de.sharewise.com/									
	05/01/13	**07/01/13**	**12/01/13**	**14/01/13**	**19/01/13**	**21/01/13**	**26/01/13**	**28/01/13**	**02/02/13**	**04/02/13**
ADIDAS AG. (DE000A1EWWW0)										
Member	11	11	11	11	12	14	14	14	15	15
Buy	4	4	4	4	6	6	6	6	5	5
Sell	7	7	7	7	8	8	8	8	10	10
Ø-TP	59.139	59.139	59.139	59.139	62.645	62.645	62.645	62.645	62.715	62.715
Analysts	16	16	16	16	16	16	16	15	15	15
Buy	9	9	9	9	9	9	9	8	8	8
Hold	6	6	6	6	6	6	6	6	6	6
Sell	1	1	1	1	1	1	1	1	1	1
Ø-TP	54.2	54.2	54.2	54.2	54.2	54.2	71.323	71.433	71.433	71.433
HEIDELBERGCEMENT (DE0006047004)										
Member	11	11	11	11	11	10	9	8	8	8
Buy	4	4	4	4	4	4	3	2	2	2
Sell	7	7	7	7	7	6	6	6	6	6
Ø-TP	40.126	40.126	40.126	40.126	40.126	40.139	39.392	37.423	37.423	37.423
Analysts	17	17	17	17	17	17	17	17	16	16
Buy	10	10	10	10	10	10	10	10	9	9
Hold	3	3	3	3	3	3	3	3	3	3
Sell	4	4	4	4	4	4	4	4	4	4
Ø-TP	42	42	42	42	42	42	47.269	47.269	47.292	47.292
RWE (DE0007037129)										
Member	14	14	12	12	10	10	11	13	13	14
Buy	9	9	8	8	8	8	10	10	10	11
Sell	5	5	4	4	2	2	2	3	3	3
Ø-TP	34.959	34.959	36.578	36.578	38.084	38.084	37.987	36.795	36.795	36.739
Analysts	22	22	22	22	22	22	22	18	18	18
Buy	6	6	6	6	5	5	5	5	5	5
Hold	9	9	9	9	10	10	10	10	10	10
Sell	7	7	7	7	7	7	7	3	3	3
Ø-TP					31	31	33.605	34.533	34.533	34.533
Siemens (DE0007236101)										
Member	12	12	12	11	10	10	11	11	12	11
Buy	11	11	9	9	9	9	8	9	9	9
Sell	76.79	76.79	79.771	79.26	78.747	78.747	80.839	80.548	81.379	81.185
Ø-TP	19	19	19	19	19	19	19	19	19	19
Analysts	4	4	4	4	4	4	4	4	4	4
Buy	13	13	13	13	13	13	13	13	13	13
Hold	2	2	2	2	2	2	2	2	2	2
Sell							83.778	83.778	83.778	83.778
Ø-TP	12	12	12	11	10	10	11	11	12	11
THYSSENKRUPP. (DE0007500001)										
Member	31	31	31	31	32	33	32	32	33	33
Buy	12	12	11	11	13	12	12	12	15	15
Sell	20	20	20	20	21	21	20	20	18	18
Ø-TP	16.075	16.075	16.389	16.389	16.857	16.762	16.754	16.754	17443	17443
Analysts	18	18	18	18	18	18	18	18	18	18
Buy	7	7	7	7	7	7	7	7	7	7
Hold	6	6	6	6	6	6	6	6	6	6
Sell	5	5	5	5	5	5	5	5	5	5
Ø-TP	16.5	16.5	16.5	16.5	16.5	16.5	18.133	18.133	18.133	18.133

Table 126. *Spekunauten Community Results (Main Experiment)*

http://www.spekunauten.de/

	17/11/12	19/11/12	24/11/12	26/11/12	01/12/12	03/12/12	08/12/12	10/12/12	15/12/12	17/12/12
ADIDAS AG. (DE000A1EWWW0)										
Members	32	32	32	32	32	32	32	32	33	33
Buy	32	32	32	32	32	32	32	32	33	33
Sell	0	0	0	0	0	0	0	0	0	0
HeldelbergCement SE (DE0006047004)										
Members	10	10	10	10	10	10	10	10	10	10
Buy	10	10	10	10	10	10	10	10	10	10
Sell	0	0	0	0	0	0	0	0	0	0
RWE (DE0007037129)										
Members	32	32	32	32	32	32	32	32	32	32
Buy	18	18	18	18	18	18	18	18	18	18
Sell	14	14	14	14	14	14	14	14	14	14
Siemens (DE0007236101)										
Members	41	41	41	41	41	41	41	41	41	41
Buy	40	40	40	40	40	40	40	40	40	40
Sell	1	1	1	1	1	1	1	1	1	1
THYSSENKRUPP. (DE0007500001)										
Members	27	27	27	27	27	27	27	27	27	27
Buy	22	22	22	22	22	22	22	22	22	22
Sell	5	5	5	5	5	5	5	5	5	5

	05/01/13	07/01/13	12/01/13	14/01/13	19/01/13	21/01/13	26/01/13	28/01/13	02/02/13	04/02/13
ADIDAS AG. (DE000A1EWWW0)										
Members	33	33	33	33	33	33	33	33	33	33
Buy	33	33	33	33	33	33	33	33	33	33
Sell	0	0	0	0	0	0	0	0	0	0
HeldelbergCement SE (DE0006047004)										
Members	10	10	10	10	10	10	10	10	10	10
Buy	10	10	10	10	10	10	10	10	10	10
Sell	0	0	0	0	0	0	0	0	0	0
RWE (DE0007037129)										
Members	32	32	32	32	32	32	32	32	32	32
Buy	18	18	18	18	18	18	18	18	18	18
Sell	14	14	14	14	14	14	14	14	14	14
Siemens (DE0007236101)										
Members	44	44	44	44	44	44	44	44	44	44
Buy	43	43	43	43	43	43	43	43	43	43
Sell	1	1	1	1	1	1	1	1	1	1
THYSSENKRUPP. (DE0007500001)										
Members	27	27	27	27	27	27	27	27	27	27
Buy	22	22	22	22	22	22	22	22	22	22
Sell	5	5	5	5	5	5	5	5	5	5

Table 127. *Stockjaeger Community Results (Main Experiment)*

http://www.stockjaeger.de/										
	17/11/12	**19/11/12**	**24/11/12**	**26/11/12**	**01/12/12**	**03/12/12**	**08/12/12**	**10/12/12**	**15/12/12**	**17/12/12**
ADIDAS AG. (DE000A1EWWW0)										
Members	2	2	2	2	2	2	2	2	2	2
Buy	2	2	2	2	2	2	2	2	2	2
Sell	0	0	0	0	0	0	0	0	0	0
Ø-Target Price	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €
HeldelbergCement AG (DE000A1EWWW0)										
Members	3	3	3	3	3	3	3	3	3	3
Buy	3	3	3	3	3	3	3	3	3	3
Sell	0	0	0	0	0	0	0	0	0	0
Ø-Target Price	53 €	53 €	53 €	53 €	53 €	53 €	53 €	53 €	53 €	53 €
RWE (DE0007037129)										
Members	2	2	2	2	2	2	2	2	2	2
Buy	1	1	1	1	1	1	1	1	1	1
Sell	1	1	1	1	1	1	1	1	1	1
Ø-Target Price	45 €	45 €	45 €	45 €	45 €	45 €	45 €	45 €	45 €	45 €
Siemens (DE0007236101)										
Members	3	3	3	3	3	3	3	3	3	3
Buy	3	3	3	3	3	3	3	3	3	3
Sell	0	0	0	0	0	0	0	0	0	0
Ø-Target Price	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €
THYSSENKRUPP. (DE0007500001)										
Members	2	2	2	2	2	2	2	2	2	2
Buy	2	2	2	2	2	2	2	2	2	2
Sell	0	0	0	0	0	0	0	0	0	0
Ø-Target Price	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €
	Only one TP availabe	Only one TP availabe	Only one TP availabe	Only one TP availabe	Only one TP availabe	Only one TP availabe	Only one TP availabe	Only one TP availabe	Only one TP availabe	Only one TP availabe
	05/01/13	**07/01/13**	**12/01/13**	**14/01/13**	**19/01/13**	**21/01/13**	**26/01/13**	**28/01/13**	**02/02/13**	**04/02/13**
ADIDAS AG. (DE000A1EWWW0)										
Members	2	2	2	2	2	2	2	2	2	2
Buy	2	2	2	2	2	2	2	2	2	2
Sell	0	0	0	0	0	0	0	0	0	0
Ø-Target Price	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €	44.5 €
HeldelbergCement AG (DE000A1EWWW0)										
Members	3	3	3	3	3	3	3	3	3	3
Buy	3	3	3	3	3	3	3	3	3	3
Sell	0	0	0	0	0	0	0	0	0	0
Ø-Target Price	53 €	53 €	53 €	53 €	53 €	53 €	53 €	53 €	53 €	53 €
RWE (DE0007037129)										
Members	2	2	2	2	2	2	2	2	2	2
Buy	1	1	1	1	1	1	1	1	1	1
Sell	1	1	1	1	1	1	1	1	1	1
Ø-Target Price	45 €	45 €	45 €	45 €	45 €	45 €	45 €	45 €	45 €	45 €
Siemens (DE0007236101)										
Members	3	3	3	3	3	3	3	3	3	3
Buy	3	3	3	3	3	3	3	3	3	3
Sell	0	0	0	0	0	0	0	0	0	0
Ø-Target Price	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €	111.67 €
THYSSENKRUPP. (DE0007500001)										
Members	2	2	2	2	2	2	2	2	2	2
Buy	2	2	2	2	2	2	2	2	2	2
Sell	0	0	0	0	0	0	0	0	0	0
Ø-Target Price	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €	37.000 €
	Only one TP available	Only one TP available	Only one TP available	Only one TP available	Only one TP available	Only one TP available	Only one TP available	Only one TP available	Only one TP available	Only one TP available

Table 128. *Stock Price Development (3-Month Period)*

Week	Delphi	Round	Share	Closing Price	Closing Price 3M later	Up / Down
1	1	1	Adidas	56.510 €	61,540 €	5,03 €
1	1	1	BAYER	60.380 €	58.210 €	-2.17 €
1	1	1	RWE	31.635 €	31.675 €	0.04 €
1	1	1	THYSSENKRUPP	21.970 €	16.615 €	-5.36 €
1	2	2	Adidas	57.220 €	61.050 €	3.83 €
1	2	2	BAYER	60.730 €	58.760 €	-1.97 €
1	2	2	RWE	32.160 €	31.810 €	-0.35 €
1	2	2	THYSSENKRUPP	21.895 €	16.390 €	-5.51 €
2	3	1	Adidas	59.850 €	60.420 €	0.57 €
2	3	1	BAYER	62.980 €	56.820 €	-6.16 €
2	3	1	RWE	33.330 €	30.220 €	-3.11 €
2	3	1	THYSSENKRUPP	20.550 €	14.830 €	-5.72 €
2	4	2	Adidas	60.340 €	59.200 €	-1.14 €
2	4	2	BAYER	64.660 €	56.940 €	-7.72 €
2	4	2	RWE	33.660 €	29.990 €	-3.67 €
2	4	2	THYSSENKRUPP	21.370 €	14.890 €	-6.48 €
3	5	1	Adidas	58.750 €	58.900 €	0.15 €
3	5	1	BAYER	64.700 €	57.800 €	-6.90 €
3	5	1	RWE	33.240 €	29.985 €	-3.26 €
3	5	1	THYSSENKRUPP	20.535 €	14.530 €	-6.01 €
3	6	2	Adidas	58.200 €	59.180 €	0.98 €
3	6	2	BAYER	65.810 €	57.050 €	-8.76 €
3	6	2	RWE	33.465 €	30.450 €	-3.02 €
3	6	2	THYSSENKRUPP	20.355 €	14.130 €	-6.23 €
4	7	1	Adidas	59.330 €	57.200 €	-2.13 €
4	7	1	BAYER	66.640 €	54.060 €	-12.58 €
4	7	1	RWE	34.730 €	28.730 €	-6.00 €
4	7	1	THYSSENKRUPP	20.380 €	12.825 €	-7.56 €
4	8	2	Adidas	59.470 €	56.280 €	-3.19 €
4	8	2	BAYER	65.650 €	53.450 €	-12.20 €
4	8	2	RWE	34.495 €	28.070 €	-6.43 €
4	8	2	THYSSENKRUPP	20.000 €	12.630 €	-7.37 €
5	9	1	Adidas	58.100 €	58.190 €	0.09 €
5	9	1	BAYER	65.220 €	55.470 €	-9.75 €
5	9	1	RWE	35.270 €	28.665 €	-6.61 €
5	9	1	THYSSENKRUPP	19.350 €	12.035 €	-7.32 €
5	10	2	Adidas	58.720 €	57.730 €	-0.99 €
5	10	2	BAYER	65.520 €	55.830 €	-9.69 €
5	10	2	RWE	35.455 €	28.760 €	-6.70 €
5	10	2	THYSSENKRUPP	19.335 €	11.640 €	-7.70 €

Table 129. *Accuracy of 3-Month Price Predictions*

E-Delphi Round	Adidas (Lay Group)	Adidas (Expert)	Adidas (Expert Group)
1	2.88%	-6.26%	-14.52%
2	2.75%	6.90%	-6.50%
3	3.53%	-5.71%	-8.50%
4	6.74%	-3.65%	-4.20%
5	5.62%	5.28%	-1.53%
6	3.64%	10.00%	-4.32%
7	9.16%	13.15%	1.91%
8	8.85%	14.66%	3.45%
9	5.47%	6.56%	-1.47%
10	2.90%	7.27%	-2.38%
Ø	5.16%	4.82%	-3.81%

E-Delphi Round	BASF (Lay Group)	BASF (Expert)	BASF (Expert Group)
1	10.01%	-5.32%	-4.76%
2	5.85%	-6.19%	4.24%
3	10.23%	0.29%	-1.83%
4	10.57%	-1.45%	5.76%
5	10.66%	15.77%	16.54%
6	13.95%	19.68%	13.60%
7	18.07%	20.92%	16.92%
8	19.62%	23.69%	18.61%
9	11.53%	20.75%	17.17%
10	16.91%	20.10%	15.01%
Ø	12.74%	10.82%	10.12%

E-Delphi Round	RWE (Lay Group)	RWE (Expert)	RWE (Expert Group)
1	2.08%	-17.94%	-0.03%
2	16.22%	-18.07%	-0.45%
3	8.58%	-12.66%	1.34%
4	15.05%	-11.85%	6.96%
5	7.03%		16.59%
6	9.37%	-7.32%	7.62%
7	28.64%	26.69%	23.81%
8	21.01%	28.79%	26.85%
9	15.44%	3.79%	17.02%
10	19.54%	3.50%	7.26%
Ø	14.30%	-0.51%	10.70%

E-Delphi Round	ThyssenKrupp (Lay Group)	ThyssenKrupp (Expert)	ThyssenKrupp (Expert Group)
1	30.23%	42.72%	41.20%
2	30.13%	43.89%	43.89%
3	36.50%	54.36%	47.87%
4	32.28%	47.31%	45.75%
5	34.59%	55.86%	48.55%
6	38.98%	53.40%	53.40%
7	40.97%	59.74%	59.74%
8	49.54%	61.85%	63.52%
9	45.30%	56.67%	61.83%
10	42.66%	69.10%	69.10%
Ø	38.12%	54.49%	53.48%

Experiment Design – Communication with Participants

This is an example of a feedback email provided after each e-Delphi run for the group members with the aggregated results of the lay group:

****

Subject: W4R2: Gruppenergebnisse Pilot-Studie "Aktien e-Delphi"

Hallo,

hier noch die Ergebnisse der zweiten e-Delphi-Runde aus Woche 4:

Adidas	**steigt**	**fällt**
Eine Woche	25%	75%
3 Monate	62.5%	37.5%
Kursziel	EUR 61.54	Average
	EUR 61,50	Median
BASF	**steigt**	**fällt**
Eine Woche	50%	50%
3 Monate	50%	50%
Kursziel	EUR 66.33	Average
	EUR 67,50	Median
RWE	**steigt**	**fällt**
Eine Woche	75%	25%
3 Monate	37.5%	62.5%
Kursziel	EUR 35,20	Average
	EUR 34,50	Median
ThyssenKrupp	**steigt**	**fällt**
Eine Woche	12.5%	87.5%
3 Monate	25%	75%
Kursziel	EUR 20.83	Average
	EUR 19,40	Median

(falls diese Tabelle bei Dir nicht richtig dargestellt wird, bitte die angehängte Grafik beachten). Gleich werde ich Dir eine Einladung mit einem direkten Link zum Aufruf der letzen Umfragerunde (Woche 5) senden.

Bei Fragen - nicht zögern: 06196 / 9997288 oder Mail an mich!

Viele Grüße
Tobias

Translation of the email to the group:

Subject: W4R2: Group results of the pilot study "Equity e-Delphi"

Hello,

here are the results from the second e-Delphi-Round of week 4:

Adidas	**up**	**down**
One Week	25%	75%
3 Month	62.5%	37.5%
Price Target	EUR 61.54	Average
	EUR 61,50	Median
BASF	**up**	**down**
One Week	50%	50%
3 Month	50%	50%
Price Target	EUR 66.33	Average
	EUR 67,50	Median
RWE	**up**	**down**
One Week	75%	25%
3 Month	37.5%	62.5%
Price Target	EUR 35,20	Average
	EUR 34,50	Median
ThyssenKrupp	**up**	**down**
One Week	12.5%	87.5%
3 Month	25%	75%
Price Target	EUR 20.83	Average
	EUR 19,40	Median

(if this table isn't displayed properly for you, please refer to the attached graphic).
I will send you an invitation with a direct link to the call for the last round of the survey (week 5) immediately.

If you have any questions don't hesitate to call: 06196 / 9997288 or email me!

Kind regards,
Tobias

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 Internetverhalten ändern, um an der Umfrage teilzunehmen?	Did you need to change your Internet usage in order to participate in 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?	Have you ever bought shares? Did you buy any 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

Interview Questionnaire (Main Experiment)

Topics and sample questions to be explored in the semi-structured interviews with all participants of the main run of the 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	Würdest Du Dich eher als emotionale oder rationale Person beschreiben?	Would you describe yourself rather as an emotional or a rational person?	General / Initial Situation
3	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
4	Bist Du bei einer Social Community wie Facebook oder Linkedin?	Are you member of a social community like facebook or linkedin?	General / Initial Situation
5	Bist Du ein Mitglied einer Aktienmarkt-Community wie sharewise oder marketocracy? Wenn nicht, könntest Du Dir vorstellen Mitglied zu werden?	Are you a member of a stock-market community like sharewise or marketocracy? If not, could you imagine becoming a member?	General / Initial Situation
6	Interessierst Du Dich für den Aktienmarkt? Wie äußert sich das?	Are you interested in the stock-market?	General / Initial Situation
7	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
8	Was war die Grundlage für Deine Entscheidungen?	How did you make your decision?	Decision-Making Process
9	Hast Du Dich auf die Umfrage-Runden vorbereitet? Wenn ja, wie?	Did you prepare for the survey rounds? If yes, how?	Decision-Making Process
10	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
11	Hast Du schon einmal selbst Aktien gekauft? Auch welche die in dem Experiment vorkommen?	Have you ever bought shares? Did you buy any used in the experiment?	Personal Impact
12	Haben/ hätten Dich die Gruppenergebnisse interessiert?	Are you (EDG)/ Would you have been (NFG) interested in the Group results?	Personal Impact
13	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
14	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
15	Du hast angegeben, dass du [XY] als Entscheidungsgrundlage genommen hast und hast einige/keine Änderungen [Anzahl der Änderungen] vorgenommen, warum? Was hat Dich dazu veranlasst?	You mentioned that your decision making is based on [XY], and you did/did not change [changes in direction] during the experiment; why? What made you do it?	Personal Impact

16	Wurdest Du von der Gruppe dabei beeinflusst worden?	Have you been influenced by the group?	Personal Impact
17	Denkst Du, dass Du neue Expertise oder Erkenntnisse hinzugewonnen hast?	Do you think you gained new expertise or knowledge during this experiment?	Personal Impact
18	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 in the survey?	Personal Impact
19	Was denkst Du über das Web-Umfrage-Tool?	What do you think about the usability of the web survey tool?	Survey Structure / Web Tool
20	Was würdest Du an der Umfrage verbessern?	What would you like be changed for the survey?	Survey Structure / Web Tool
21	Du hast [x mal] kein Kursziel angeben, warum?	You didn't enter a price target [in X of the cases], why?	Survey Structure / Web Tool
22	Denkst Du es ist leichter einen konkreten Kurs (in Euro) als Kursziel zu nennen oder die prozentuale Veränderung?	Do you think its more easy to enter a concrete price (in Euro) as price target or a percentage change?	Survey Structure / Web Tool
23	Hast Du weitere Kommentare oder Anregungen?	Any further comments or suggestions?	General Issues

Sample Translation of an Interview

The following is a translation of an interview transcription. All interviews have been conducted in German language and prepared with a semi-structured interview questionnaire (see also Interview Questionnaire (Main Experiment) on page 296). The participant who gave this interview is neither a particularly good, nor an especially bad predictor:

Interviewer: Did you feel it was easy to answer the questions?

Participant: To estimate the prices?

Interviewer: Yes.

Participant: Well, it is difficult. I think in these cases ten times further. What is intended by the question? I would say it is simple and clear. As a non-expert, it is of course difficult, because I have no idea. And because of that it’s a question of how much time I want to invest in researching in order to be in the picture, and how far I just say do it off the top of my head and follow the group. And since I'm not an expert, how should I know it clearly. The questions were simple and clear, and it was repeated continuously, to that extent, the process was simple, but assessing stocks is difficult for me because I have no idea of this topic.

Interviewer: Would you describe yourself rather as an emotional or a rational person?

Participant: Emotional. I try to follow a rational course, however, I think that my emotions play a major role. I try to get along rationally, but I assume that my emotions drive me more than I would like.

Interviewer: Did you need to change your Internet usage in order to participate in the survey?

Participant: No.

Interviewer: Are you member of a social community like facebook or linkedin?

Participant: Yes, I'm a member of facebook, but I don't really like it. I just kind of slipped into it. Well, it was quite nice for getting in contact with a few people abroad ten or seven years ago, but I don't have it on my iPhone or use it regularly, I look at it now and again, more in a passive way.

Interviewer: Are you a member of a stock-market community like sharewise or marketocracy?

Participant: No.

Interviewer: Could you imagine becoming a member?

Participant: No, I think . . . we did finance at school and a bit at university too, and I do find it a fascinating subject, but there are much more interesting things, especially when I don't absolutely have to concern myself with it. So . . . perhaps if you just know the bare bones, a general understanding, but not really for its own sake. I don't think it will ever happen that I take an interest in it again.

Interviewer: Are you interested in the stock-market?

Participant: Hmm. Interested, well I think it's important to have an overall understanding of how the world wags, and that includes our money-system. I find the share system totally complex and actually a bit crazy, and I don't know if I should condone that, but good, I don't have much of a clue about it and that's why I'm careful about making any statements. I'm more interested in the overall picture than in the details – or the idea that someday I might say I must invest in shares.

Interviewer: How would you self-assess your knowledge about the stock market? (1, no knowledge; 10, expert)

Participant: Perhaps 3, but I could . . . with the things I've now learned, I ought to know more, but when you're not involved with it and when it's some time ago, it's quite difficult to say anything about it. So I have more of a feeling: Oh dear, I've forgotten everything. So between 2 and 3, but I think my knowledge will come back quite quickly. When you've studied something the basic knowledge is there

Interviewer: How did you make your decisions?

Participant: Well, it's always the sort of thing, what do I know about it, and then it was more or less off the top of my head. And then I looked at some pages on the Internet – to be honest I don't really know how serious they are, and if everything's right and how much you can believe it. On the other hand I also found that it's really difficult to interpret these things. I no longer know about all that. Yes, I kept looking at it, but then the group, obviously it gave me a push, although I have no idea who's in the group, whether they're all people like me who have no idea then . . . of course as a result I can't really rely on them

Interviewer: Did you prepare for the survey rounds?

Participant: Honestly, I have to admit, No. I sometimes intended to ask someone and take a closer interest in it, but then everything again went so fast, fast

Interviewer: Did you use external sources for the experiment?

Participant: Well, I googled it. What was it? Finanzen.net it was mostly. I just googled share prices and looked to see how they had been in the past, and of course at the end of the day you have to make an assessment. And, well, to really know anything I would have had to analyse the companies and much more with the data. I mean, on the Internet you can find some of that, more the numbers for the company, but that's as far as I get, after that I find it difficult to say anything about it, so then I just relied on general opinions that were available.

Interviewer: Have you ever bought shares?

Participant: No.

Interviewer: Are you interested in the Group results?

Participant: Yes, but by then I couldn't remember – or couldn't always remember – which box I'd ticked. Then I always thought, I ought to print it out so that I could see it in progress, how it changed every time and what I had chosen. That is, it would have been interesting but it only helped me to a limited extent, but the overall picture would have been helpful.

Interviewer: Do you think the e-Delphi-experiment / the group results influenced your decisions?

Participant: Yes, yes, because I needed something from somewhere, because I was there on the Internet, and looked to see what info I could find, [and] I was relatively uncertain. So then again the group was a help and a guideline for me

Interviewer: You changed your decision a couple of times in round 2; why? [The interviewer shares with the participant a hard-copy with the changes in recommendation]

Participant: Well, I probably looked at the group again. I actually did the same thing every time – except 1, 2 or 3 I probably had no time and relied more on the group. Otherwise I always went to the Internet and took a quick look, and then at what the group was saying, and then no longer remembered what I'd said last time.

Interviewer: You mentioned that your decision-making is based on external sources, like news, why? What made you do it?

Participant: That was probably more out of laziness, though I actually had the impression that that helped me most with my laziness and to get something out of it for my decision, -- if that really was the right website. No guarantee. I tried to give a decision that I found relatively acceptable in a very short time, although I was really uncertain, so in an ideal situation I'd have got an expert to explain it all to me, and then ask him what he recommended and act accordingly

Interviewer: Do you think you gained new expertise or knowledge during this experiment?

Participant: Yes, sometimes I did have the impression that I was similar to the group . But I can not say for sure.

Interviewer: With or against?

Participant: Damn. Really? Well, sometime in the middle of all this I realised that I had interpreted something incorrectly again. I can remember – certainly on one occasion – it will obviously have been the same for the others – when I differed from the group, where I was on the Internet and thought, No, I'd do that differently, and if I didn't know who was in the group, how expert are they experts I can rely on and how many of them are non-experts like me? To that extent I was uncertain. Though I did mostly look to see what the group was saying and I'm sure I sometimes interpreted it wrongly.

Interviewer: Do you think you gained new expertise or knowledge during this experiment?

Participant: No, but I should again have prepared myself better and simply talked to an expert again.

Interviewer: Do you care more about news now, in particular news of the companies in the survey?

Participant: No.

Interviewer: It's not that you're going to be more attentive when news about Adidas or ThyssenKrupp is on the radio?

Participant: I did during the survey, of course you think about how the company is doing now and how things are developing, and I do believe that if I'd heard it on the news I'd have pricked up my ears, but somehow I can't now remember any occasion when it reminded me of the study. .

Interviewer: What do you think about the usability of the web survey tool?

Participant: In principle, I don't understand the whole big picture and for that reason it is difficult. It is ultimately a simple tool that it is easy to use, simple from the practical point of view. How far that leads to correct, good results I obviously can't tell. I do believe that if you collaborate with a group it strengthens your power of decision, but for that you'd really need more background information about who is in the group and in general.

Interviewer: What would you like be changed for the survey?

Participant: Well, I believe, what kept on coming, but I don't know what your objective is, for me personally it would have been better to be better prepared, so some kind of introduction with an explanation. But the question obviously is if it is your aim for me to think about the topic and then refresh my knowledge to be able to answer the questions better. That is the question. But my feeling is that I would have preferred to have a better introduction and information.

Interviewer: Some other issues that you would like to mention? Perhaps the timing and content?

Participant: No, I found the content easy because of the presentation and how it was run and implemented. My problem was simply the knowledge itself, to be able to make the assessment. No, otherwise I thought it was good. The distribution of the email and the link as well. Then there was only the fact that I have Monday off, so then the fact that it was on a Monday was a bit stupid.

Interviewer: You didn't enter a price target in 3 of the cases, why?

Participant: Oh, probably I missed it.

Interviewer: Any reason for this?

Participant: No, then I forgot, I thought I couldn't click on. Then I must have been half asleep.

Interviewer: Do you think its more easy to enter a concrete price—in Euro—as price target or a percentage change?

Participant: Yes, obviously. Well, I had a real problem with percentages. Even now I don't know if I really . . . so I didn't check it at first, Then I looked to see what the others in the group were doing. But it's true, I would have found it easier.

Interviewer: Do you have any further comments for me?

Participant: Good question,but I’ve actually said all the important things. No.

Comparative Analysis of Best and Worst Predictors

Table 130. *Comparative Analysis of Best and Worst Predictors (Code Matrix)*

	204	503	511	516	101	604	617	507	520	2	515	521
Rank of Participant (Based on Overall Predictive Accuracy)	1	2	3	4	5	6	54	55	56	57	58	59
Efforts of participants and comments regarding participation												
Stock Market is my Hobby					1							
It was interesting/fun to participate						1						
Experiment different to my usual investment decision-making				1								
Over time more intense or more systematic				1			1					
In the beginning more intense efforts		1		1		1		1		1		
Mood and feelings of participants												
Wish and reality differ quite a lot										1		
I don't know if I was correct												1
I was not very good										1		
Typical behaviour							1					
Decision approach and techniques												
Mean Reversion					1							
Market Environment (Sentiment)	1				1							
Short-Term Momentum					1							
Group as contra-indicator					1	1						
Interview_Questions												
Q1												
Easy to understand, but difficult to answer			1	1	1			1		1	1	
Easy to answer	1	1				1	1		1			1
Q2												
Emotional		1	1					1	1		1	
Rational	1			1	1	1	1			1		1
I try to be rational/Emotional is not good						1						
Q3												
No, I'm always or often online		1		1	1	1	1	1	1	1	1	
No, maybe not online but mobile.	1											
Yes, usually I don't check daily			1									1
Q4												
yes	1	1	1				1	1	1	1	1	
yes, but not very active				1	1	1						
no												1
Q5												
Maybe, it could be interesting		1	1		1	1		1				
No, not for me.	1			1			1		1	1	1	1

Not, anymore												
Q6												
Not or very little interested in the stock market		1									1	2
Is interested in stock market	1		1	1	1	1	1	1	1	1		
Q8												
Economy, Politics and Macro						1		1				
Business Cycles/Seasons						1					1	1
News and Media		1					1	1	1	1	1	2
Technical Analysis		1	1	1								
Company Information or Analysis	1		1		1							
Market sentiment	1			1	2					2		
Opinions of others (incl. analysts)	1			1								
Peers, Friends, Experts					1							
Intuition/Gut feeling/Smart Guess			1		1		1		1	1		
Q9												
Yes			1									
Yes, but just a little or sometimes		1		1				1		1	1	
No	1					1	1		1			1
Q10												
No							1	1				
no, but...												1
Yes, Newspapers, Magazines						1				1		
Maybe speaking with experts	1											
Yes, Internet Portals		1	1	1		1			1	1	1	
I checked some facts and figures					1							
Q11												
No			1						1			1
Yes	1	1		1	1	1	1	1		1	1	
Yes\No, not more easy.							1	1				
Yes\Yes, more easy	1			1						1		
Q12												
No		1										
Yes	1		1	1		1	1	1		1	1	1
Not often					1							
Q13												
No	1	1			1	2	1		1	2		
No\I hope not.					1							
No\I would listen to experts, but not the group.	1											
No\No or no, I don't think so		1				2	1		1	2		
Yes			1	1							1	1
Yes\Yes, I think so.				1							1	1
Yes\Confirmation or reassurance of their			1									

own opinion												
I don't know			1									
Q14/15												
Different Mood/Weekend more time						1	1					
Company specific issues						1	1					
Group communication/group influence										1		
Market sentiment / Political Issues			3									
Experts	1		1	1								
Intuition/Gut Feeling		2		1				1		1	1	
No cautious decisions/by chance				1						1	1	
Technical Analysis	1		1									
Newsflow		1	1				1	1				2
Q17												
Yes						1					1	
Maybe a little			2	1						2		
No, I don't think so.	1	1			1		1	1	1			1
Q18												
No, I don't care or don't pay attention									1			
No, I had already a strong background before.					1							
Yes, I think so.	1	1	2	1		1	1	1			1	
Q22												
Wouldn't ask any of these questions	1											
Euro more easy									1	1	1	
% is better		1	1	1		1	1	1				1
Doesn't matter to me.					1							

Annex III

Preference for Intuition and Deliberation (PID): An Inventory for Assessing Affect- and Cognition-Based Decision-Making

Instructions: Please answer all the following questions about your life in general. Your answers should correspond to the way you generally make decisions. Circle the number that best represents your opinion. 1 means that you very much disagree; 5 means that you very much agree.

Question	German Version	English Translation	Preference
1	Bevor ich Entscheidungen treffe, denke ich meistens erst mal gründlich nach.	Before making decisions I first think them through.	Preference for deliberation
2	Ich beobachte sorgfältig meine innersten Gefühle.	I listen carefully to my deepest feelings.	Preference for intuition
3	Bevor ich Entscheidungen treffe, denke ich meisten erst mal über meine Ziele nach, die ich erreichen will.	Before making decisions I usually think about the goals I want to achieve.	Preference for deliberation
4	Bei den meisten Entscheidungen ist es sinnvoll sich ganz auf sein Gefühl zu verlassen.	With most decisions it makes sense to completely rely on your feelings.	Preference for intuition
5	Ich mag Situationen nicht, in denen ich mich auf meine Intuition verlassen muss.	I don’t like situations that require me to rely on my intuition.	Preference for intuition
6	Ich denke über mich nach.	I think about myself.	Preference for deliberation
7	Ich schmiede lieber ausgefeilte Pläne, als etwas dem Zufall zu überlassen.	I prefer making detailed plans rather than leaving things to chance.	Preference for deliberation
8	Ich ziehe Schlussfolgerungen lieber aufgrund meiner Gefühle, Menschenkenntnis und Lebenserfahrung.	I prefer drawing conclusions based on my feelings, my knowledge of human nature, and my experience of life.	Preference for intuition
9	Bei meinen Entscheidungen spielen Gefühle eine große Rolle.	My feelings play an important role in my decisions.	Preference for intuition
10	Ich bin perfektionistisch.	I am a perfectionist.	Preference for deliberation
11	Wenn ich meine Entscheidungen rechtfertigen muss, denke ich vorher besonders gründlich nach.	I think about a decision particularly carefully if I have to justify it.	Preference for deliberation
12	Wenn es darum geht, ob ich anderen vertrauen soll, entscheide ich aus dem Bauch heraus.	When it comes to trusting people, I can usually rely on my gut feelings.	Preference for intuition
13	Ich nehme bei einem Problem erst mal die Daten und Fakten auseinander, bevor ich mich entscheide.	When I have a problem I first analyze the facts and details before I decide.	Preference for deliberation
14	Ich denke erst nach bevor ich handle.	I think before I act.	Preference for deliberation
15	Ich mag lieber gefühlsbetonte	I prefer emotional people.	Preference for

	Personen.		intuition
16	Ich denke über meine Ziele und Pläne stärker nach als andere Personen.	I think more about my plans and goals than other people do.	Preference for deliberation
17	Ich bin ein sehr intuitiver Mensch.	I am a very intuitive person.	Preference for intuition
18	Ich mag emotionale Situationen, Diskussionen und Filme.	I like emotional situations, discussions, and movies.	Preference for intuition

Note. Inventory for Assessing Affect- and Cognition-Based Decision-Making. Adapted from "Präferenz für Intuition und Deliberation (PID) [Preference for Intuition and Deliberation (PID): An Inventory for Assessing Affect- and Cognition-Based Decision-Making]," by C. Betsch, 2004, p. 183.

Annex IV

Data Generated with the Main Experiment: Analysis and Data Sets

Table 131. *Comparison of 1-Week Predictions Grouped by PID Scale Score*

	Participants	Correct	Wrong	Sum	Percentage of correct answers
PID-D	22	838	870	1708	49,1%
PID-I	11	430	425	855	50,3%
PID-S minus	17	688	667	1355	50,8%
PID-S plus	9	328	327	655	50,1%

Table 132. *Comparison of 1-Month Predictions Grouped by PID Scale Score*

	Participants	Correct	Wrong	Sum	Percentage of correct answers
PID-D	22	851	857	1708	49,8%
PID-I	11	407	448	855	47,6%
PID-S minus	17	672	683	1355	49,6%
PID-S plus	9	352	303	655	53,7%

Table 133. *Comparison of 3-Month Predictions Grouped by PID Scale Score*

	Participants	Correct	Wrong	Sum	Percentage of correct answers
PID-D	22	921	787	1708	53,9%
PID-I	11	483	372	855	56,5%
PID-S minus	17	738	617	1355	54,5%
PID-S plus	9	402	253	655	61,4%

Table 134. *Predictions (1-Week) with steady upward or downward trend and without trend*

1-Week Predictions										
Trend Intact						No Trend				
Participant ID	Correct	Wrong	Excluded	Correct (%)		Participant ID	Correct	Wrong	Excluded	Correct (%)
101	146	104	35	58.4%		101	64	86	15	42.7%
102	79	111	95	41.6%		102	51	59	55	46.4%
103	69	81	135	46.0%		103	36	39	90	48.0%
104	77	167	41	31.6%		104	58	88	19	39.7%
105	113	107	65	51.4%		105	72	58	35	55.4%
14	89	66	130	57.4%		14	51	44	70	53.7%
15	119	141	25	45.8%		15	66	74	25	47.1%

2	142	143		49.8%		2	88	77		53.3%
201	122	98	65	55.5%		201	68	62	35	52.3%
202	131	154		46.0%		202	89	76		53.9%
203	160	105	20	60.4%		203	100	60	5	62.5%
204	178	97	10	64.7%		204	92	58	15	61.3%
205	132	118	35	52.8%		205	73	77	15	48.7%
34	100	110	75	47.6%		34	70	70	25	50.0%
36	57	73	155	43.8%		36	43	52	70	45.3%
38	151	134		53.0%		38	84	81		50.9%
4	52	38	195	57.8%		4	18	17	130	51.4%
501	203	82		71.2%		501	107	58		64.8%
502	54	26	205	67.5%		502	26	19	120	57.8%
503	163	122		57.2%		503	87	78		52.7%
504	93	102	90	47.7%		504	52	53	60	49.5%
505	48	47	190	50.5%		505	27	28	110	49.1%
506	62	163	60	27.6%		506	38	87	40	30.4%
507	44	56	185	44.0%		507	21	29	115	42.0%
508	83	57	145	59.3%		508	47	38	80	55.3%
509	106	84	95	55.8%		509	44	41	80	51.8%
510	120	115	50	51.1%		510	70	70	25	50.0%
511	156	114	15	57.8%		511	79	76	10	51.0%
512	117	93	75	55.7%		512	68	47	50	59.1%
513	173	112		60.7%		513	87	78		52.7%
514	186	99		65.3%		514	84	81		50.9%
515	30	70	185	30.0%		515	15	35	115	30.0%
516	166	114	5	59.3%		516	79	66	20	54.5%
517	149	136		52.3%		517	81	84		49.1%
518	120	100	65	54.5%		518	70	60	35	53.8%
519	152	133		53.3%		519	88	77		53.3%
520			285						165	
521	119	146	20	44.9%		521	71	89	5	44.4%
601	91	89	105	50.6%		601	59	61	45	49.2%
602	109	146	30	42.7%		602	56	64	45	46.7%
603	27	18	240	60.0%		603	18	12	135	60.0%
604	140	110	35	56.0%		604	85	65	15	56.7%
605	128	87	70	59.5%		605	82	53	30	60.7%
606	96	139	50	40.9%		606	64	76	25	45.7%
607	139	146		48.8%		607	81	84		49.1%
608	110	175		38.6%		608	70	95		42.4%
609	97	118	70	45.1%		609	68	67	30	50.4%
610	90	160	35	36.0%		610	60	90	15	40.0%
611	133	152		46.7%		611	77	88		46.7%
612	122	138	25	46.9%		612	88	77		53.3%
613	129	96	60	57.3%		613	66	59	40	52.8%
614	125	130	30	49.0%		614	65	80	20	44.8%
615	138	147		48.4%		615	77	88		46.7%
616	107	153	25	41.2%		616	68	72	25	48.6%
617	133	152		46.7%		617	82	83		49.7%
618	129	136	20	48.7%		618	86	74	5	53.8%

619	14	36	235	28.0%		619	6	19	140	24.0%
620	81	104	100	43.8%		620	39	51	75	43.3%
621	110	145	30	43.1%		621	65	80	20	44.8%
Sum	6509	6395	3911	50.12%		Sum	3726	3710	2299	49.56%

Table 135. *Predictions (1-Month) with steady upward or downward trend and without trend*

1-Month Predictions										
Trend Intact						No Trend				
Participant ID	Correct	Wrong	Excluded	Correct (%)		Participant ID	Correct	Wrong	Excluded	Correct (%)
101	215	125	45	63.2%		101	30	30	5	50.0%
102	166	104	115	61.5%		102	19	11	35	63.3%
103	119	86	180	58.0%		103	11	9	45	55.0%
104	169	161	55	51.2%		104	31	29	5	51.7%
105	125	165	95	43.1%		105	30	30	5	50.0%
14	100	110	175	47.6%		14	15	25	25	37.5%
15	172	168	45	50.6%		15	28	32	5	46.7%
2	141	244		36.6%		2	29	36		44.6%
201	175	115	95	60.3%		201	40	20	5	66.7%
202	176	209		45.7%		202	39	26		60.0%
203	167	193	25	46.4%		203	33	32		50.8%
204	255	105	25	70.8%		204	45	20		69.2%
205	179	166	40	51.9%		205	26	29	10	47.3%
34	133	172	80	43.6%		34	17	28	20	37.8%
36	111	79	195	58.4%		36	24	11	30	68.6%
38	194	191		50.4%		38	31	34		47.7%
4	51	64	270	44.3%		4	4	6	55	40.0%
501	190	195		49.4%		501	35	30		53.8%
502	55	45	285	55.0%		502	10	15	40	40.0%
503	268	117		69.6%		503	47	18		72.3%
504	82	168	135	32.8%		504	18	32	15	36.0%
505	61	69	255	46.9%		505	9	11	45	45.0%
506	150	150	85	50.0%		506	25	25	15	50.0%
507	63	72	250	46.7%		507	7	8	50	46.7%
508	121	74	190	62.1%		508	19	11	35	63.3%
509	119	136	130	46.7%		509	11	9	45	55.0%
510	185	130	70	58.7%		510	30	30	5	50.0%
511	230	135	20	63.0%		511	30	30	5	50.0%
512	163	132	90	55.3%		512	17	13	35	56.7%
513	143	242		37.1%		513	22	43		33.8%
514	206	179		53.5%		514	29	36		44.6%
515	57	73	255	43.8%		515	8	12	45	40.0%
516	243	132	10	64.8%		516	32	18	15	64.0%
517	222	163		57.7%		517	33	32		50.8%
518	169	141	75	54.5%		518	21	19	25	52.5%
519	200	185		51.9%		519	35	30		53.8%
520			385						65	
521	110	250	25	30.6%		521	25	40		38.5%

601	166	84	135	66.4%		601	34	16	15	68.0%
602	146	184	55	44.2%		602	24	21	20	53.3%
603	38	12	335	76.0%		603	22	3	40	88.0%
604	173	172	40	50.1%		604	32	23	10	58.2%
605	140	165	80	45.9%		605	20	25	20	44.4%
606	145	170	70	46.0%		606	35	25	5	58.3%
607	143	242		37.1%		607	27	38		41.5%
608	164	221		42.6%		608	26	39		40.0%
609	84	211	90	28.5%		609	26	29	10	47.3%
610	145	195	45	42.6%		610	30	30	5	50.0%
611	161	224		41.8%		611	29	36		44.6%
612	171	189	25	47.5%		612	34	31		52.3%
613	179	116	90	60.7%		613	36	19	10	65.5%
614	150	190	45	44.1%		614	30	30	5	50.0%
615	195	190		50.6%		615	30	35		46.2%
616	113	222	50	33.7%		616	22	43		33.8%
617	168	217		43.6%		617	32	33		49.2%
618	192	168	25	53.3%		618	43	22		66.2%
619	24	46	315	34.3%		619	1	4	60	20.0%
620	178	72	135	71.2%		620	17	8	40	68.0%
621	159	176	50	47.5%		621	26	39		40.0%
Sum	8719	8711	5285	50.38%		Sum	1491	1419	925	51.18%

Table 136. *Predictions (3-Month) with steady upward or downward trend and without trend*

3-Month Predictions										
Trend intact						No trend				
Participant ID	Correct	Wrong	Excluded	Correct (%)		Participant ID	Correct	Wrong	Excluded	Correct (%)
101	236	114	50	67.4%		101	34	16		68.0%
102	183	87	130	67.8%		102	17	13	20	56.7%
103	87	118	195	42.4%		103	8	12	30	40.0%
104	201	139	60	59.1%		104	24	26		48.0%
105	123	182	95	40.3%		105	22	23	5	48.9%
14	87	133	180	39.5%		14	13	17	20	43.3%
15	164	186	50	46.9%		15	31	19		62.0%
2	113	287		28.3%		2	12	38		24.0%
201	152	148	100	50.7%		201	28	22		56.0%
202	210	190		52.5%		202	35	15		70.0%
203	134	246	20	35.3%		203	16	29	5	35.6%
204	284	91	25	75.7%		204	36	14		72.0%
205	152	198	50	43.4%		205	33	17		66.0%
34	169	141	90	54.5%		34	16	24	10	40.0%
36	125	80	195	61.0%		36	10	10	30	50.0%
38	211	189		52.8%		38	24	26		48.0%
4	60	65	275	48.0%		4			50	
501	146	254		36.5%		501	24	26		48.0%
502	66	49	285	57.4%		502	4	6	40	40.0%
503	296	104		74.0%		503	44	6		88.0%

504	157	108	135	59.2%		504	23	12	15	65.7%
505	59	76	265	43.7%		505	6	9	35	40.0%
506	150	160	90	48.4%		506	30	10	10	75.0%
507	48	82	270	36.9%		507	12	8	30	60.0%
508	105	85	210	55.3%		508	20	15	15	57.1%
509	78	167	155	31.8%		509	12	18	20	40.0%
510	237	93	70	71.8%		510	28	17	5	62.2%
511	280	105	15	72.7%		511	20	20	10	50.0%
512	179	111	110	61.7%		512	26	9	15	74.3%
513	192	208		48.0%		513	33	17		66.0%
514	222	178		55.5%		514	28	22		56.0%
515	64	76	260	45.7%		515	6	4	40	60.0%
516	247	138	15	64.2%		516	28	12	10	70.0%
517	254	146		63.5%		517	26	24		52.0%
518	193	122	85	61.3%		518	22	13	15	62.9%
519	236	164		59.0%		519	24	26		48.0%
520			400			520			50	
521	112	268	20	29.5%		521	28	17	5	62.2%
601	166	99	135	62.6%		601	14	21	15	40.0%
602	187	158	55	54.2%		602	23	7	20	76.7%
603	39	21	340	65.0%		603	11	4	35	73.3%
604	264	86	50	75.4%		604	36	14		72.0%
605	136	164	100	45.3%		605	24	26		48.0%
606	212	118	70	64.2%		606	33	12	5	73.3%
607	268	132		67.0%		607	37	13		74.0%
608	266	134		66.5%		608	34	16		68.0%
609	186	129	85	59.0%		609	19	16	15	54.3%
610	207	143	50	59.1%		610	33	17		66.0%
611	189	211		47.3%		611	26	24		52.0%
612	179	196	25	47.7%		612	31	19		62.0%
613	165	135	100	55.0%		613	35	15		70.0%
614	245	105	50	70.0%		614	30	20		60.0%
615	301	99		75.3%		615	34	16		68.0%
616	259	91	50	74.0%		616	36	14		72.0%
617	154	246		38.5%		617	26	24		52.0%
618	232	148	20	61.1%		618	28	17	5	62.2%
619	30	45	325	40.0%		619			50	
620	165	90	145	64.7%		620	10	10	30	50.0%
621	263	92	45	74.1%		621	32	13	5	71.1%
Sum	10125	7930	5545	55.31%		Sum	1355	930	665	58.41%

Data Generated by the Main Experiment: Use of Information Sources and Different Decision-Making Principles/Influences

Each subscript letter indicates a subset of correct 1-week categories whose column proportions do not differ significantly from each other on the .05 level.

Table 137. *Crosstabs 1-Week Predictions / Decision Making Basis*

			1 Week		
			wrong	correct	Sum
Financial ratios (SQ001)	no	Frequency	2073_a	2078_a	4151
		Expected Frequency	2075.5	2075.5	4151.0
		% in 1-Week Predictions	91.2%	91.5%	91.4%
	yes	Frequency	199_a	194_a	393
		Expected Frequency	196.5	196.5	393.0
		% in 1-Week Predictions	8.8%	8.5%	8.6%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	$.070^a$	1	.792
Frequency of valid Data Sets	4544		

			1-Week Predictions		
			wrong	correct	Sum
Fundamental analysis (SQ002)	no	Frequency	2198_a	2171_b	4369
		Expected Frequency	2184.5	2184.5	4369.0
		% in 1-Week Predictions	96.7%	95.6%	96.1%
	yes	Frequency	74_a	101_b	175
		Expected Frequency	87.5	87.5	175.0
		% in 1-Week Predictions	3.3%	4.4%	3.9%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	4.333^a	1	.037
Frequency of valid Data Sets	4544		

			1-Week Predictions		
			wrong	correct	Sum
Group results (SQ003)	no	Frequency	2070_a	2110_b	4180
		Expected Frequency	2090.0	2090.0	4180.0
		% in 1-Week Predictions	91.1%	92.9%	92.0%
	yes	Frequency	202_a	162_b	364
		Expected Frequency	182.0	182.0	364.0
		% in 1-Week Predictions	8.9%	7.1%	8.0%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	4.778[a]	1	.029
Frequency of valid Data Sets	4544		

			1-Week Predictions		
			wrong	correct	Sum
Company (SQ004)	no	Frequency	1859_a	1872_a	3731
		Expected Frequency	1865.5	1865.5	3731.0
		% in 1-Week Predictions	81.8%	82.4%	82.1%
	yes	Frequency	413_a	400_a	813
		Expected Frequency	406.5	406.5	813.0
		% in 1-Week Predictions	18.2%	17.6%	17.9%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	.253[a]	1	.615
Frequency of valid Data Sets	4544		

			1Week Predictions		
			wrong	correct	Sum
Intuition (SQ005)	no	Frequency	337_a	363_a	700
		Expected Frequency	350.0	350.0	700.0
		% in 1-Week Predictions	14.8%	16.0%	15.4%
	yes	Frequency	1935_a	1909_a	3844
		Expected Frequency	1922.0	1922.0	3844.0
		% in 1-Week Predictions	85.2%	84.0%	84.6%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	1.142[a]	1	.285
Frequency of valid Data Sets	4544		

			1-Week Predictions		
			wrong	correct	Sum
Market Sentiment (SQ006)	no	Frequency	1223_a	1294_b	2517
		Expected Frequency	1258.5	1258.5	2517.0
		% in 1-Week Predictions	53.8%	57.0%	55.4%
	yes	Frequency	1049_a	978_b	2027
		Expected Frequency	1013.5	1013.5	2027.0
		% in 1-Week Predictions	46.2%	43.0%	44.6%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	4.490[a]	1	.034
Frequency of valid Data Sets	4544		

			1-Week Predictions		
			wrong	correct	Sum
News (SQ008)	no	Frequency	1545[a]	1590[a]	3135
		Expected Frequency	1567.5	1567.5	3135.0
		% in 1-Week Predictions	68.0%	70.0%	69.0%
	yes	Frequency	727[a]	682[a]	1409
		Expected Frequency	704.5	704.5	1409.0
		% in 1-Week Predictions	32.0%	30.0%	31.0%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	2.083[a]	1	.149
Frequency of valid Data Sets	4544		

			1-Week Predictions		
			wrong	correct	Sum
Expert opinions	no	Frequency	2039[a]	2000[a]	4039
(SQ007)		Expected Frequency	2019.5	2019.5	4039.0
		% in 1-Week Predictions	89.7%	88.0%	88.9%
	yes	Frequency	233[a]	272[a]	505
		Expected Frequency	252.5	252.5	505.0
		% in 1-Week Predictions	10.3%	12.0%	11.1%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	3.388[a]	1	.066
Frequency of valid Data Sets	4544		

			1-Week Predictions		
			wrong	correct	Sum
Technical analysis	no	Frequency	1928[a]	1958[a]	3886
(SQ009)		Expected Frequency	1943.0	1943.0	3886.0
		% in 1-Week Predictions	84.9%	86.2%	85.5%
	yes	Frequency	344[a]	314[a]	658
		Expected Frequency	329.0	329.0	658.0
		% in 1-Week Predictions	15.1%	13.8%	14.5%
Sum		Frequency	2272	2272	4544
		Expected Frequency	2272.0	2272.0	4544.0
		% in 1-Week Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	1.599[a]	1	.206
Frequency of valid Data Sets	4544		

Each subscript letter indicates a subset of correct 1-month categories whose column proportions do not differ significantly from each other on the .05 level.

Table 138. *Crosstabs 1-Month Predictions / Decision Making Basis*

			1-Month Predictions		
			wrong	correct	Sum
Financial ratios (SQ001)	no	Frequency	2097_a	2054_b	4151
		Expected Frequency	2077.3	2073.7	4151.0
		% in 1-Month Predictions	92.2%	90.5%	91.4%
	yes	Frequency	177_a	216_b	393
		Expected Frequency	196.7	196.3	393.0
		% in 1-Month Predictions	7.8%	9.5%	8.6%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	4.312[a]	1	.038
Frequency of valid Data Sets	4544		

			1Month Predictions		
			wrong	correct	Sum
Fundamental analysis (SQ002)	no	Frequency	2196_a	2173_a	4369
		Expected Frequency	2186.4	2182.6	4369.0
		% in 1-Month Predictions	96.6%	95.7%	96.1%
	yes	Frequency	78_a	97_a	175
		Expected Frequency	87.6	87.4	175.0
		% in 1Month Predictions	3.4%	4.3%	3.9%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	2.180[a]	1	.140
Frequency of valid Data Sets	4544		

			1-Month Predictions		
			wrong	correct	Sum
Group results (SQ003)	no	Frequency	2096_a	2084_a	4180
		Expected Frequency	2091.8	2088.2	4180.0
		% in 1-Month Predictions	92.2%	91.8%	92.0%
	yes	Frequency	178_a	186_a	364
		Expected Frequency	182.2	181.8	364.0
		% in 1-Month Predictions	7.8%	8.2%	8.0%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	.207[a]	1	.649
Frequency of valid Data Sets	4544		

			1-Month Predictions		Sum
			wrong	correct	
Company (SQ004)	no	Frequency	1857_a	1874_a	3731
		Expected Frequency	1867.1	1863.9	3731.0
		% in 1-Month Predictions	81.7%	82.6%	82.1%
	yes	Frequency	417_a	396_a	813
		Expected Frequency	406.9	406.1	813.0
		% in 1-Month Predictions	18.3%	17.4%	17.9%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	.616[a]	1	.432
Frequency of valid Data Sets	4544		

			1-Month Predictions		Sum
			wrong	correct	
Intuition (SQ005)	no	Frequency	337_a	363_a	700
		Expected Frequency	350.3	349.7	700.0
		% in 1-Month Predictions	14.8%	16.0%	15.4%
	yes	Frequency	1937_a	1907_a	3844
		Expected Frequency	1923.7	1920.3	3844.0
		% in 1-Month Predictions	85.2%	84.0%	84.6%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	1.196[a]	1	.274
Frequency of valid Data Sets	4544		

			1-Month Predictions		Sum
			wrong	correct	
Market Sentiment (SQ006)	no	Frequency	1230_a	1287_a	2517
		Expected Frequency	1259.6	1257.4	2517.0
		% in 1-Month Predictions	54.1%	56.7%	55.4%
	yes	Frequency	1044_a	983_a	2027
		Expected Frequency	1014.4	1012.6	2027.0
		% in 1-Month Predictions	45.9%	43.3%	44.6%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	3.123[a]	1	.077
Frequency of valid Data Sets	4544		

			1-Month Predictions		
			wrong	correct	Sum
News (SQ008)	no	Frequency	1585_a	1550_a	3135
		Expected Frequency	1568.9	1566.1	3135.0
		% in 1-Month Predictions	69.7%	68.3%	69.0%
	yes	Frequency	689_a	720_a	1409
		Expected Frequency	705.1	703.9	1409.0
		% in 1-Month Predictions	30.3%	31.7%	31.0%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	1.069[a]	1	.301
Frequency of valid Data Sets	4544		

			1-Month Predictions		
			wrong	correct	Sum
Expert opinions (SQ007)	no	Frequency	2050_a	1989_b	4039
		Expected Frequency	2021.3	2017.7	4039.0
		% in 1-Month Predictions	90.1%	87.6%	88.9%
	yes	Frequency	224_a	281_b	505
		Expected Frequency	252.7	252.3	505.0
		% in 1-Month Predictions	9.9%	12.4%	11.1%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	7.351[a]	1	.007
Frequency of valid Data Sets	4544		

			1-Month Predictions		
			wrong	correct	Sum
Technical analysis (SQ009)	no	Frequency	1995_a	1891_b	3886
		Expected Frequency	1944.7	1941.3	3886.0
		% in 1-Month Predictions	87.7%	83.3%	85.5%
	yes	Frequency	279_a	379_b	658
		Expected Frequency	329.3	328.7	658.0
		% in 1-Month Predictions	12.3%	16.7%	14.5%
Sum		Frequency	2274	2270	4544
		Expected Frequency	2274.0	2270.0	4544.0
		% in 1-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	17.977[a]	1	.000
Frequency of valid Data Sets	4544		

Each subscript letter indicates a subset of correct 3-month categories whose column proportions do not differ significantly from each other on the .05 level.

Table 139. *Crosstabs 3-Month Predictions / Decision Making Basis*

			3-Month Predictions		
			wrong	correct	Sum
Financial ratios (SQ001)	no	Frequency	1833 [a]	2318 [a]	4151
		Expected Frequency	1839.8	2311.2	4151.0
		% in 3-Month Predictions	91.0%	91.6%	91.4%
	yes	Frequency	181 [a]	212 [a]	393
		Expected Frequency	174.2	218.8	393.0
		% in 3-Month Predictions	9.0%	8.4%	8.6%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	.524 [a]	1	.469
Frequency of valid Data Sets	4544		

			3-Month Predictions		
			wrong	correct	Sum
Fundamental analysis (SQ002)	no	Frequency	1927 [a]	2442 [a]	4369
		Expected Frequency	1936.4	2432.6	4369.0
		% in 3-Month Predictions	95.7%	96.5%	96.1%
	yes	Frequency	87 [a]	88 [a]	175
		Expected Frequency	77.6	97.4	175.0
		% in 3-Month Predictions	4.3%	3.5%	3.9%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	2.144 [a]	1	.143
Frequency of valid Data Sets	4544		

			3-Month Predictions		
			wrong	correct	Sum
Group results (SQ003)	no	Frequency	1882 [a]	2298 [b]	4180
		Expected Frequency	1852.7	2327.3	4180.0
		% in 3-Month Predictions	93.4%	90.8%	92.0%
	yes	Frequency	132 [a]	232 [b]	364
		Expected Frequency	161.3	202.7	364.0
		% in 3-Month Predictions	6.6%	9.2%	8.0%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	10.413 [a]	1	.001
Frequency of valid Data Sets	4544		

			3-Month Predictions		
			wrong	correct	Sum
Company (SQ004)	no	Frequency	1633_a	2098_a	3731
		Expected Frequency	1653.7	2077.3	3731.0
		% in 3-Month Predictions	81.1%	82.9%	82.1%
	yes	Frequency	381_a	432_a	813
		Expected Frequency	360.3	452.7	813.0
		% in 3-Month Predictions	18.9%	17.1%	17.9%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	2.591[a]	1	.107
Frequency of valid Data Sets	4544		

			3-Month Predictions		
			wrong	correct	Sum
Intuition (SQ005)	no	Frequency	270_a	430_b	700
		Expected Frequency	310.3	389.7	700.0
		% in 3-Month Predictions	13.4%	17.0%	15.4%
	yes	Frequency	1744_a	2100_b	3844
		Expected Frequency	1703.7	2140.3	3844.0
		% in 3-Month Predictions	86.6%	83.0%	84.6%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	11.089[a]	1	.001
Frequency of valid Data Sets	4544		

			3-Month Predictions		
			wrong	correct	Sum
Market Sentiment (SQ006)	no	Frequency	1059_a	1458_b	2517
		Expected Frequency	1115.6	1401.4	2517.0
		% in 3-Month Predictions	52.6%	57.6%	55.4%
	yes	Frequency	955_a	1072_b	2027
		Expected Frequency	898.4	1128.6	2027.0
		% in 3-Month Predictions	47.4%	42.4%	44.6%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	11.558[a]	1	.001
Frequency of valid Data Sets	4544		

			3 Month Predictions		
			wrong	correct	Sum
News (SQ008)	no	Frequency	1388_a	1747_a	3135
		Expected Frequency	1389.5	1745.5	3135.0
		% in 3Month Predictions	68.9%	69.1%	69.0%
	yes	Frequency	626_a	783_a	1409
		Expected Frequency	624.5	784.5	1409.0
		% in 3-Month Predictions	31.1%	30.9%	31.0%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	.009[a]	1	.923
Frequency of valid Data Sets	4544		

			3-Month Predictions		
			wrong	correct	Sum
Expert opinions (SQ007)	no	Frequency	1820_a	2219_b	4039
		Expected Frequency	1790.2	2248.8	4039.0
		% in 3-Month Predictions	90.4%	87.7%	88.9%
	yes	Frequency	194_a	311_b	505
		Expected Frequency	223.8	281.2	505.0
		% in 3-Month Predictions	9.6%	12.3%	11.1%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	8.031[a]	1	.005
Frequency of valid Data Sets	4544		

			3-Month Predictions		
			wrong	correct	Sum
Technical analysis (SQ009)	no	Frequency	1745_a	2141_a	3886
		Expected Frequency	1722.4	2163.6	3886.0
		% in 3-Month Predictions	86.6%	84.6%	85.5%
	yes	Frequency	269_a	389_a	658
		Expected Frequency	291.6	366.4	658.0
		% in 3-Month Predictions	13.4%	15.4%	14.5%
Sum		Frequency	2014	2530	4544
		Expected Frequency	2014.0	2530.0	4544.0
		% in 3-Month Predictions	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2-sided)
Pearson-Chi-Squared	3.691[a]	1	.055

Annex V

Quantitative Factor Analysis

The following correlation matrix shows that the variable "Commitment" (COMSQ001) and the remaining variables correlate only at a relatively low level, with the exception of the variable Skill Self-Assessment. The variables PID-D and PID-I are negatively correlated with each other (-0.222). A relatively high, in each case negative, correlation of -0.563 indicates the variable PID-I with "Emotionality/Rationality Self-Assessment" (Emo. Self-Assessment) and of -0.384 with Skill Self-Assessment. Skill Self-Assessment is positively correlated with "Emotionality/Rationality Self-Assessment" (Emo. Self-Assessment) ($r = 0.378$). Almost all correlations are significant, see table 140 below.

Table 140. *Factor Analysis – Correlation Matrix*

Correlation Matrix		COMSQ001	PID-D	PID-I	Emo. Self-Assessment	Skill Self-Assessment	PAALL
Correlation	COMSQ001	1.000	-.019	.084	-.043	.232	-.132
	PID-D	-.019	1.000	-.222	.098	.141	.196
	PID-I	.084	-.222	1.000	-.563	-.384	.041
	Emo. Self-Assessment	-.043	.098	-.563	1.000	.378	-.134
	Skill Self-Assessment	.232	.141	-.384	.378	1.000	.032
	PAALL	-.132	.196	.041	-.134	.032	1.000
Sig. (1-sided)	COMSQ001		.105	.000	.002	.000	.000
	PID-D	.105		.000	.000	.000	.000
	PID-I	.000	.000		.000	.000	.003
	Emo. Self-Assessment	.002	.000	.000		.000	.000
	Skill Self-Assessment	.000	.000	.000	.000		.015
	PAALL	.000	.000	.003	.000	.015	

a. Determinant = .432

The explained total variance was, for the first two factors, approximately 53% (see Table 141). Since the underlying variables can certainly not be measured without error, the factor analysis was examined using the extraction method “principal axis factor analysis” (and not the principal component analysis, which presupposes error-free variables).

Table 141. *Factor Analysis – Total Variance Explained*

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings[a]
	Total	% of Variance	Cumulative %	Total
1	1.967	32.778	32.778	1.452
2	1.257	20.948	53.725	.658
3	1.090	18.171	71.897	
4	.771	12.843	84.739	
5	.496	8.267	93.006	
6	.420	6.994	100.000	

Extraction Method: Principal Axis Factoring.

a. When components are correlated, sums of squared loadings cannot be added to obtain a total value.

The following scree plot (see Figure 33) shows that with one, or a maximum of two, components the total variance can be explained (the first two factors together explain approximately 53%).

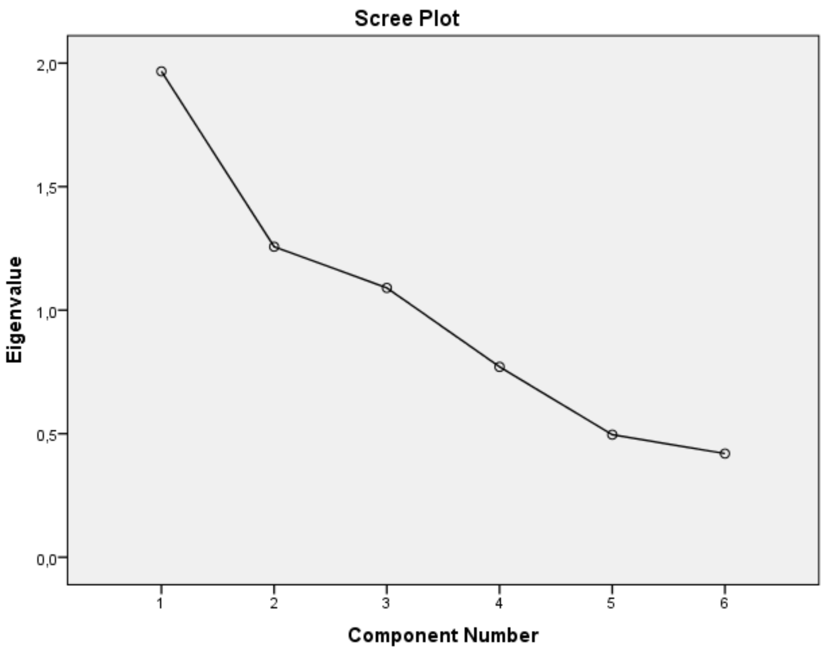


Figure 33: Scree Plot

The following pattern matrix (see Table 142) reflects the loading of the variables on the two components: On component 1, the variables PID-I, Emo. Self-Assessment and Skill Self-Assessment (loading higher than 0.51), while the second component loads, in particular, the variable PAALL. The two components are only slightly correlated (r = 0.058), due to the rotation method Oblimin. Since the variables include all personality traits, the independence of factors is certain (this would allow the rotation method Varimax).

Table 142. *Factor Analysis – Pattern Matrix*[a]

Variable	Component	
	1	2
Preference for Intuition (PID-I)	-.783	-.020
Emo. Self-Assessment	.726	-.146
Skill Self-Assessment	.510	.010
Overall Predictive Accuracy (PAALL)	-.080	.732
Preference for Deliberation (PID-D)	.229	.287
Commitment (COMSQ001)	.020	-.148

Extraction Method: Principal Axis Factor Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
a. Rotation converged in 2 iterations.

Table 143. *Factor Analysis – Component Correlation Matrix*

Component	1	2
1	1.000	.058
2	.058	1.000

Extraction Method: Principal Axis Factor Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

The following section elaborates on a factor analysis with an alternative rotation method (Varimax). The analysis was conducted analogously to the factor analysis with the Oblimin rotation method.

Table 144. *Total Variance Explained (Varimax Rotation Method)*

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.967	32.778	32.778	1.454	24.226	24.226
2	1.257	20.948	53.725	.658	10.966	35.192
3	1.090	18.171	71.897			
4	.771	12.843	84.739			
5	.496	8.267	93.006			
6	.420	6.994	100.000			

Extraction Method: Principal Axis Factor Analysis.

The following scree plot analogously shows (see Figure 34) that with one, maximally two components the total variance is explained (the first two factors together explain approximately 53%).

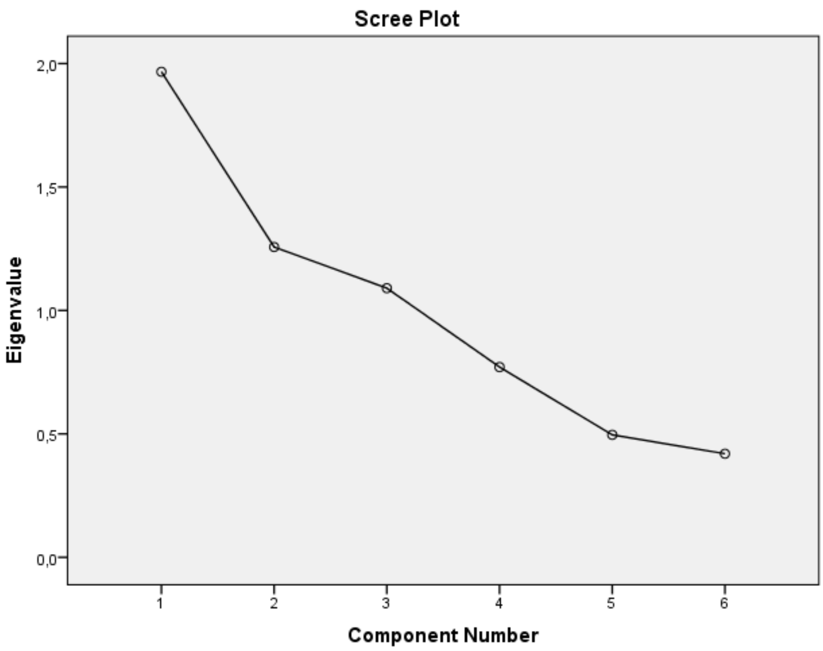


Figure 34: Scree Plot (with Varimax Rotation Method**)**

The following pattern matrix (see Table 145) reflects the loading of the variables on the two components: On component 1, the variable PID-I, Emo. Self-Assessment and Skill Self-Assessment (loading higher than 0.51), while the second component loads, in particular, the variable PAALL. These are very similar values to the Oblimin rotation method.

Table 145. *Rotated Component Matrix (Varimax Rotation Method)*

Variable	Component	
	1	2
Preference for Intuition (PID-I)	-.783	-.044
Emo. Self-Assessment	.722	-.123
Skill Self-Assessment	.510	.026
Overall Predictive Accuracy (PAALL)	-.060	.729
Preference for Deliberation (PID-D)	.236	.294
Commitment (COMSQ001)	.016	-.148

Extraction Method: Principal Axis Factor Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

The graph below (see Figure 35) shows the position of the variables spanned in two components: On component 1, the variables PID-I (negative), Emo. Self-Assessment and Skill Self-Assessment (both positive) show high values, while the variable (PAALL) shows high values on component 2. The variable COMSQ001 is relatively insignificant (values close to 0 on both components): This was also determined using the Oblimin rotation method.

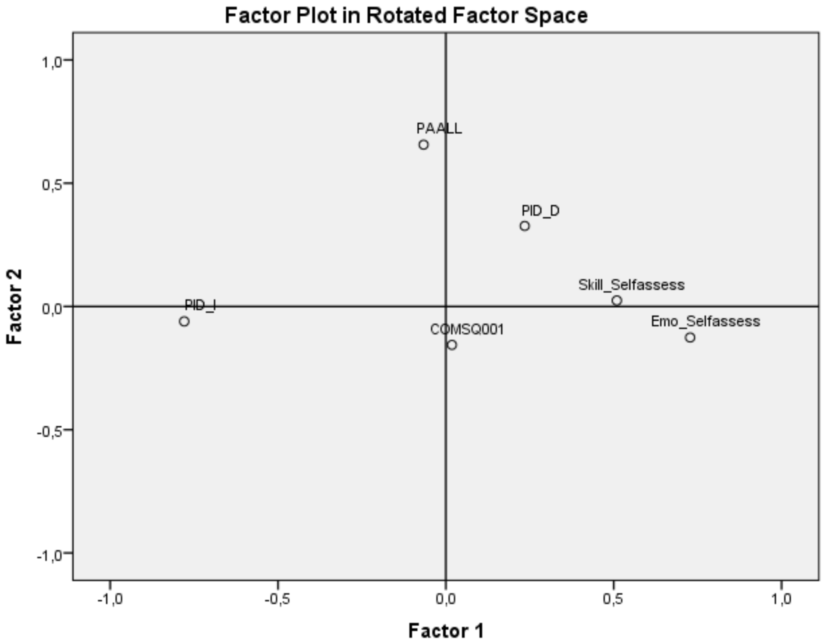


Figure 35: Component Plot in Rotated Space (Varimax Rotation Method)

The following coefficient matrix of component scores is based on the above graph (see Figure 35) The values are very similar to the values using the Oblimin method: differences arise only in the second or third decimal place.

Table 146. *Component Score Coefficient Matrix* (Varimax Rotation Method)

	Component	
	1	2
Commitment (COMSQ001)	.038	-.050
Preference for Deliberation (PID-D)	.068	.149
Preference for Intuition (PID-I)	-.498	-.093
Emo. Self-Assessment	.376	-.099
Skill Self-Assessment	.159	-.004
Overall Predictive Accuracy (PAALL)	-.003	.684

Extraction Method: Principal Axis Factor Analysis.
Rotation Method: Varimax with Kaiser Normalization.