UNIVERSITY OF
GLOUCESTERSHIRE

This is a peer-reviewed, final published version of the following document and is licensed under Creative Commons: Attribution 3.0 license:

Endress, Tobias (2018) "Deliberated Intuition" in Stock Price Forecasting. Economics and Sociology, 11 (3). pp. 11-27. doi:10.14254/2071-789X.2018/11-3/1

Official URL: http://www.economics-sociology.eu/?596,en_deliberated-intuition-in-stock-price-forecasting
DOI: 10.14254/2071-789X.2018/11-3/1
EPrint URI: https://eprints.glos.ac.uk/id/eprint/6041

Disclaimer

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.

ECONOMICS & Sociology

Endress, T. (2018). "Deliberated Intuition" in Stock Price Forecasting. *Economics and Sociology*, *11*(3), 11-27. doi:10.14254/2071-789X.2018/11-3/1

"DELIBERATED INTUITION" IN STOCK PRICE FORECASTING

Tobias Endress,
University of Gloucestershire, Cheltenham, United Kingdom,
E-mail: tobias@endress.info

Received: May, 2018
1st Revision: May, 2018
Accepted: June, 2018

DOI: 10.14254/2071-789X.2018/11-3/1

ABSTRACT. Financial analysis is a topic of interest for both academic research and businesses. Financial analysts are important elements of economic interactions. Nevertheless, there are doubts about the quality of their predictions. Special crowdsourcing platforms facilitate group decisions as an alternative to traditional financial analysis. The objective of this paper is to investigate the quality of predictions by individuals and groups using this alternative approach. Various groups – consisting of laypeople but also financial professionals – were formed purposefully to generate equity forecasts. The data from the experiment suggest that some variables, in terms of participants' characteristics, have a significant impact on the quality of predictions. The results show that intuition plays an important role in the decision-making process. Also, good predictors base their intuition on several factors. The results led to an explanatory model, which we introduce as "deliberated intuition", a practice process being different for each individual. It appears that thinking about the problem in different ways and with various techniques contribute to making good predictions. The model may help in designing teams for traditional financial analysts.

JEL Classification: D70, D91, G11, G17, N24.

Keywords: forecasting, decision-making, deliberated intuition, equity predictions, stock prices.

Introduction

Financial analysis is an interesting subject for both academic research and business purposes. The work of financial analysts is of special significance for brokers and investment banks and almost every financial newspaper, stock market journal or TV programme that deals with financial topics (Stanzel, 2007). Nevertheless, there are many doubts about the quality of work provided by these experts. The Internet, particularly crowdsourcing platforms, facilitate group decision-making as an alternative approach to traditional financial analysis. The present study investigates this alternative approach and elaborates upon a variation of a crowdsourced group decision-making process as suggested by Kaplan (2001). The overall aim is to explore, analyse and compare the quality of equity predictions by individuals and groups using the Internet in order to build a theory of this process.

The aims of this paper is to contribute with domain-specific insights about the underlying key mechanisms in the context of equity predictions. The research questions are: What are the underlying key mechanisms that influence the predictive process, and how might this process be improved? The purpose of this study is to provide insights on the group

decision-making process for Internet communities which focus on stock-trading based on predicting share prices. The intent of this two-phase, sequential mixed-methods empirical study is to develop an explanatory model of the group decision-making process within the Internet communities.

1. Literature review

Some authors claim that forecasting "is a skill that can be cultivated" (Tetlock & Gardner, 2015, p. 4). This might be particularly relevant for professional financial analysts and their forecasts of share prices. However, a myriad of studies have already been conducted to assess the quality of forecasts by financial analysts (Aiolfi, Rodriguez, & Timmermann, 2009; Bolliger, 2004; Clement, 1999; Fleischer, 2005; Lin, 2017; Roger, Roger, & Schatt, 2015; Stanzel, 2007). In general these studies show little evidence that it is possible to generate accurate predictions that in the long run, after transaction costs, create profits higher than the market average (Malkiel, 2007; Stanzel, 2007). Additionally it is difficult to present and communicate predictions with adequate precision (Løhre & Teigen, 2017).

Some authors argue that there is a body of empirical and theoretical evidence indicating an advantage in combining forecasts (Armstrong, 2001; Hong & Page, 2004; Silver, 2012). The modality of crowdsourcing (Prpić & Shukla, 2016) applied by special interest online communities which inspired this experiment can be seen as open collaboration (Prpić, Taeihagh, & Melton, 2015). Aside from the aggregation algorithm, expertise and cognition were found to influence the prediction process (Budescu & Chen, 2015; Jackson, Kleitman, Stankov, & Howie, 2017; Steyvers & Miller, 2015). Numerous authors are working to enhance the concept of crowd wisdom (e.g., Davis-Stober, Budescu, Broomell, & Dana, 2015; Endress & Gear, 2018; Olsson & Loveday, 2015; Prelec, Seung, & McCoy, 2017; Prpić & Shukla, 2016). Numerous factors with influence on individual decision making have been identified (e.g., Jackson et al., 2017; Nguyen, 2018; Tetlock & Gardner, 2015). However, there is limited literature about price predictions based on a crowdsourced approach (e.g., Chen et al., 2018; Jame et al., 2016, Kaplan 2001).

Many factors impact rational choice, decision-making and forecasting. Individual characteristics, particularly cognitive abilities, are described as a predictor of an individual's decision competence (Jackson et al., 2017). However, there is limited knowledge about the characteristics that impact financial (share price) predictions. Models such as prospect theory (Kahneman & Tversky, 1979) are frequently cited in behavioural finance (Barberis & Thaler, 2003; Camerer, Loewenstein, & Rabin, 2011; Dhami, 2016; Paddock et al., 2015). Herbert Simon's model of bounded rationality may be able to provide more understanding in the context of financial markets, for which incomplete information is inherent to the subject matter. Simon (1955, 1956) introduced an enhanced model, the behavioral model of rational choice, as an alternative to the "economic man" and utility theory. Simon's model takes into consideration that access to information and the computational capacities of man are limited, and some authors have worked to enhance his model (e.g., Kasthurirathna & Piraveenan, 2015; Marwala, 2013).

Sometimes having more information may be counter-productive to making a 'right decision'. Gigerenzer (2008) demonstrated this phenomenon by identifying that in some cases, simple heuristics (e.g., 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) may lead to good decisions with less information. This means that more knowledge and information being available does not necessarily lead to better decisions – a finding that could be especially useful when examining investment decisions, for which probably nobody has complete information. But precise quantitative information also has 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).

It appears helpful to keep in mind that forecasting *"is a skill that can be cultivated"* (Tetlock & Gardner, 2015, p. 4). However, facilitating its cultivation is still not easy. Literature suggests that researchers should focus on the underlying role of intuition in the process (Glöckner & Witteman, 2010). Domain knowledge might be an important factor in the decision process (Dane, Rockmann, & Pratt, 2012). It might be the case that *"complex models often give more precise (but not necessarily more accurate) answers, [as] they can trip a forecaster's sense of overconfidence"* (Silver, 2012, p. 225). It remains difficult for forecasters to establish realistic beliefs about what they actually know (Parker & Stone, 2014).

In summary, the literature suggests a lack of reliable mechanisms to identify and assess investment processes. Literature on traditional approaches to equity research highlights the weaknesses and generally describes the quality of the predictions as problematic. Kahneman (2003) points out that a *"general framework . . . is not a substitute for domain-specific concepts and theories"* (p. 717). The aim of this research is to contribute with domain-specific insights in the context of equity predictions based on a crowdsourcing approach.

2. Pilot of Method

A pilot test of the study's method was conducted with a small group (11 participants) and three experienced financial analysts to benchmark the group over five e-Delphi cycles (five weeks).

The field experiment was conducted following an e-Delphi approach (Dalkey & Helmer-Hirschberg, 1962; Lindqvist & Nordänger, 2007). Shares were selected from four companies in different business sectors. Each participant in the pilot was asked to provide predictions of the movement (up or down) of every share for one-week and three-month periods as well as stock price predictions for a three-month period. The data for analysis were gathered from the experiment and from subsequent interviews.

The combination of quantitative and qualitative approaches and the triangulation of both was aimed 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).

The pilot identified 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 decision-making types. The results also indicated ways to improve the survey design and to adjust the structure and process. The pilot experiment demonstrated the feasibility of the experiment and showed that the tools and setup are capable of addressing the aim of the study. The pilot also provided valuable insights into improving the planned research approach—in particular, the e-Delphi method.

2.1. Adjustment of group design and feedback loop

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

Another finding of the pilot was that people did not change their predictions very often after receiving group feedback through the e-Delphi method. The literature suggests more changes and a stronger convergence of the group's decision (Dalkey, 1969; Dalkey & Helmer-Hirschberg, 1962). We decided 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. Accordingly, an interesting experiment might 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 rounds 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 was possible to determine the effect of the feedback on the group's decision-making more clearly.

2.2. Participant assessment

To understand more about the group decision-making process it is helpful to understand more about the decision-making processes of the individual group participants. In order to do so, all participants in the main experiment were individually assessed for their age, gender, education level, profession and decision-making type. While the questions about age, gender, education level and profession were quite easy to answer, the question about the decision-making type was not. Betsch (2004) developed an approach to addressing this issue (Schunk & Betsch, 2006; Traufetter, 2009) by creating and thoroughly testing a questionnaire to determine people's preference for intuition and/or deliberation (PID score). The participants' types were assessed to understand the reasons for particular predictions.

2.3. Enhancements to the online questionnaire

The analysis of the pilot's procedure and results 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 few participants provided information about their decision-making process in the free-text field on the online form. Nevertheless, the participant interviews during the pilot run indicated some clusters of different types and sources for the decision-making process. In order to simplify the answer options and to get more information, these types were provided as tick-box fields of the participants' share estimate group, so that they might 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 predicting an actual price target for the three-month period; accordingly, it might be a good idea to change the question from a concrete stock price to a price movement as a percentage for this period. Additionally, this question was changed into an optional question, in case anyone was uncomfortable with answering it. Another change was to introduce a question about their level of confidence in their predictions (from "not at all" to "absolutely sure", on a 1-5 scale).

2.4. Participant interviews

All of the participants in the pilot were interviewed. The questions were intended to gain a deeper understanding of the participants' decision-making process and improve the design of the planned experiment. All of the participants agreed that the questions were easy to understand, and they all felt able to estimate or at least guess whether the stock price was going up or down.

The interviews of the indicated different bases for the individual predictions. *Table 1* shows a summary of the different answers to '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 were grouped into nine clusters of decision-making influences.

Table 1. Clusters of 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 first round)
Intuition	Gut feeling, intuition, instinct, guess
Market sentiment	General market situation and market outlook
News	Including daily press, Internet, and business and finance news
Technical analysis	Index development, price-movement, momentum, etc.

These clusters were transferred into easy-to-understand options for the lay participants in the main experiment. The participants ticked a box on the online survey for each weekly prediction of a share and could add a comment.

3. Main Experiment

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

3.1. Method

The design and approach of the main experiment were in principle similar to the design of the pilot run (Endress, 2015). There were just a few changes, including an enhanced online questionnaire, more interview questions and participants being asked to enter the target price, not as the total amount in euros but as a change in percent. Deciding on the appropriate sample size for research is not easy. 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 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, these are now very old data and an old analysis in a different contextual task. Since was also necessary to keep in touch with all of the participants over three months and to allow for a moderate drop-out quota, an initial number for the panel of about 10 participants (N=10) for the panel appeared 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, more than one lay group was included in the main experiment to compare survey designs and measure the differences. In addition to determining 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 of the participants were purposefully selected from the researcher's personal and professional networks. A key criterion for selection was ensuring the diversification of the group in terms of age, gender, education level and professional background, etc.

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 (10 weeks with a two-week 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 five participants
- Professional investors group (PG), with a group size of five participants
- e-Delphi-group (EDG), with a group size of 21 lay participants
- Interactive group (IG), with a group size of seven lay participants
- Non-feedback group (NFG), with a group size of 21 lay participants

In addition, the single expert/ financial analyst estimations were analysed as individual expert judgements within their 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 certain groups (EDG, IG, PG); two groups did not receive any feedback from their group members (AG, NFG). In a second round, the participants could provide different responses.

The shares included five companies in 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 total, the main experiment was set up to gather up to 17,700 individual judgements about equity predictions (i.e., 5,900 individual judgements about equity predictions for each period).

Each participant in the experiment was asked to estimate the movement of every share (up or down) over a one-week, a one-month and a three-month period as well as predict a stock price change in percent for a three-month period. Group results with an undecided voting result, i.e. the 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)—these undecided rounds were excluded from the analysis. The approach was tested with a pilot experiment prior to the main run. The combination of quantitative and qualitative approaches and the triangulation of both promised to create a more holistic understanding of the decision-making process by 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). Mixed methods are appropriate when the research questions consist of confirmatory and explanatory questions (Teddlie & Tashakkori, 2009). 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 been an established key to creating professional knowledge for many years (Schön, 1983). Balton (2010) stated 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 creation and enhancement of appropriate methods for

developing an explanatory schema and for starting to build theory. The qualitative analysis was aided by MAXQDA, a software tool for qualitative and mixed-methods data analysis.

3.2. Results

Quantitative data analysis was conducted in a sequential approach. The first step was a univariate analysis. In the second step, a multi-criteria analysis and data reduction techniques were applied. Both approaches are aimed at providing inform an understanding of factors that influence the decision-making process and forecast quality. A discriminant analysis identified some variables with potential impacts on decision-making and predictive accuracy. *Table 2* provides an overview of several significant variables for the different prediction periods.

No single variable reached a significance level in all of the periods. However, a few variables reached significance in different periods e.g., "commitment" (COMSQ001), "group results" (SQ003), "market sentiment" (SQ006). However, not all of the variables turned out to be significant. Also, some variables appeared to be comparatively unimportant in this context, since they did not reach significance level for any period, in particular, "use of company information" (SQ004), "news" (SQ008), and "preference for intuition" (PID-I) (see *Table 2*).

Table 2. Overview of Variables Based on the Discriminant Analysis

	Variables with significant differences between correct and wrong predictions		
	1-week predictions	1-month predictions	3-month predictions
Commitment (COMSQ001)	X	X	
Financial ratios (SQ001)		X	
Fundamental analysis (SQ002)	X		
Group results (SQ003)	X		X
Company (SQ004)			
Intuition (SQ005)			X
Market sentiment (SQ006)	X		X
News (SQ008)			
Expert opinions (SQ007)		X	X
Technical analysis (SQ009)		X	
Age group	X		
PID-D		X	X
PID-I			
Group no	X		
Univ. degree	X	X	

Source: own compilation.

3.3. Qualitative Interview Analysis

Overall, 59 people participated in the main experiment. Among them, 25 participants were interviewed face to face, 29 participants were interviewed by telephone and three participants were interviewed 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 interviews were semi-structured and aided by a questionnaire to direct the course of the discussion.

The interview analysis added valuable context information to the online survey data. The analysis of the in-depth qualitative interviews delivered insight into the participants'

decision-making processes and revealed that financial predictions were a rather emotional issue for many of the participants. One example was the answer to question 8—"How did you make your decision?"—to which most of the participants referred to their intuition and gut feeling. Typical statements included "Well, my gut feeling" and "Mostly intuition". The professionals also 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 was not the same for professionals and some laypeople. When asked to describe intuition and gut feeling, they answered, "it 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"; "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".

For the answers to question 15, a few topics were mentioned by all of the participants, e.g., "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 best and worst predictors reveals some differences. None of the worst six predictors relied on expert opinions, but three of the top six predictors mentioned expert opinions as an influence on, or basis for, their decisions. The use of technical analysis seemed similar: it was not mentioned by the worst six predictors, but at least two of the top six predictors mentioned that their decisions were at least partially based on technical analysis. Generally, it appeared that the top predictors tended to question the reliability of their sources more than the weak predictors and that they were more willing to adapt their decision-making approach, or at least question the source of information. As one of the best predictors put it, "I have to know the expert to get a picture and to trust his opinion". The poor predictors rarely mentioned any substantive discussion about quality of their sources or basic principles *(Table 3*).

Table 3. Comparative Analysis of Best and Worst Predictors (Code Matrix)

	Participant Interview Analysis (Frequency of Answers)											
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 interested or have very little interest in the stock market		1									1	2
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 six worst predictors revealed that all of the participants understood the questions and that they were equally difficult for good and poor predictors to answer (three of each mentioned that it was difficult). There was also no perceivable difference in their use of the Internet. All of the participants used the Internet on at least a daily basis.

There were differences between the top and worst predictors in other domains. It seems that for the good predictors, the ideal was 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 were interested in the stock market, while at least two of the poor predictors mentioned that they were not or were only a little interested in this topic. A particularly striking difference was in the bases on which they made decisions: while poor predictors mentioned that they relied on news (and their gut feeling) to a large extent, the picture for the top predictors was much more differentiated (this is particularly reflected in the answers to questions 8 and 14/15). Top predictors also emphasized slightly more that they were more sensitive to news related to the companies in the experiment (Q18).

Another difference between top and poor predictors was that none of the top predictors thought that it was easier to make concrete predictions in euros than predict 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.

Are more rational people better at financial predictions than more intuitive people?

It might be conceivable that rational people have an advantage in financial decision-making, such as predictions of stock price movements. Two variables in the data set might reflect the participants' rationality. First is the self-assessment by the participants taken from the interview data, which was coded with the categories "emotional", "rather emotional", "rather rational", and "rational". The second variable is the so-called PID scale value.

The analysis of the PID scale was more informative (Endress & Gear, 2015). In order to measure this value, all of the 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) (Betsch, 2004; Schunk & Betsch, 2006). The PID scale provides four categories with which to group people according to their preferences. There are people who prefer deliberative decisions (PID-D) or intuitive decisions (PID-I), and people with situationally varying preferences, whereas people can either have a preference for both intuitive and deliberate decision-making strategies (PID-S plus) or lack a preference for any strategy (PID-S minus). The PID-D group was most represented (N=22), followed by PID-S minus (N=17).

Following the idea that rational people have an advantage, rational people would be expected to 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 4. 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 who preferred deliberative decision-making was 50.9% compared to 51.5% for people who preferred intuitive decision-making (see *Table 4*), which is not a significant difference. The direct comparison of PID-D's and PID-I's prediction quality tested using a chi-square test results in a chi-square=0.189 and *p-value*=0.66. A comparison that included all PID types, in particular the preference for both intuitive and deliberate decision-making strategies (PID-S plus), led to significantly better predictions, which might be more interesting. The direct comparison of all four categories' predictive quality with a chi-square test resulted in a chi-square=10.084 and *p-value*=0.018. These results are, to some extent, in agreement with Tetlock's metaphor of *the hedgehog and the fox*. In his analysis, the aggregated success rate of the foxes' predictions was considerably better compared with that of the 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.

The data support the suggestion that intuitive people are better at financial decision-making compared with rational people, although the data were not significant in the 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 for the PID-S plus participants apparently had significantly higher accuracy.

Do educational level and gender impact the ability to predict stock prices?

Education is a very important component that influences people's decision-making strategies. It might be a sensible assumption that level of education influences decision-making ability and predictive accuracy. Accordingly, it might be assumed that the participants' educational level had a major impact on their ability to predict stock prices. The data gathered in the experiment did—at least, to some extent—conform with this assumption. Even so, there was only a weakly significant difference in the predictive accuracy of people with versus without university degrees (chi-square=2.967, *p-value*=0.084). People without university degrees were, with 50.4% correct answers, slightly less often correct than people with university degrees, who produced just 52.1% correct answers (see *Table 5*).

Table 5. 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%

Following the stereotype that women 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 than male participants – at least, for short-term predictions. Since all of the expert participants in this experiment were male, an analysis of the lay people only is sensible for assessing this assumption (see *Table 6*). A direct comparison of the predictions by the female and male participants in the experiment shows that the female participants were indeed able to provide predictions with significantly higher accuracy (chi-square=14.681, *p-value*<0.001). Additional tests and more data might be helpful to clarify the relationship and effect.

Table 6. Comparison of All Lay people Predictions, Grouped by Gender

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

4. Discussion

The experimental method provided a rich data set with qualitative and quantitative components. The analysis of these data sets provided valuable insights into the decision-making processes 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.

In terms of the participants' individual characteristics several variables had significant impacts on the quality of equity predictions: educational level, gender and PID-score. Some

relevant variables were also related to a particular prediction, like confidence and decision-making approach. While no single factor makes a good predictor, many factors can influence the quality of equity predictions. Still, it needs to be noted that there are noise and uncertainty in the decision-making process and that not all variables have the same significance. Flexibility and scrutiny of sources and approaches appear to be keys for good investment decisions.

A combination of approaches appears to be a superior strategy. In particular, this can be documented by the PID score analysis. Higher significance was observed for the direct comparison of all four categories. The predictions of PID-S-plus participants apparently had significantly higher accuracy. Still, these findings are from an experiment with a limited number of participants that should be repeated with a larger sample size and in different settings. Further research would be helpful to gain a better understanding of the possible limitations and underlying mechanisms.

The findings of this experiment suggest that people who prefer both intuition and deliberation (type PID-S-plus) had 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 in the hiring process for equity analysts. The experiment's findings not only underline the fact that "intuition is an important component of professional competence in the domain of the stock market" (Harteis & Gruber, 2008, p. 83) but also accentuate 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 improving of investment processes. Further research and additional data might be helpful to confirm the findings and allow inference for a wider range of settings and conditions.

Deliberated intuition model

The research findings suggest a theory of the prediction process based on "deliberated intuition" defined here as a considered decision to adopt an intuitive approach to predict share movements after reaching a limit for rational analysis. The factors influencing this considered decision are shown in *Figure 1*, comprising the individuals' personality traits, their 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 prediction quality. This model of the 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 one's social context, training and experience. The model of deliberated intuition proposed here views prediction as a practice process which will be different for each individual. It is not a dual-processing model with two modes of process such as rational deliberation or intuitive processes (e.g., Tversky & Kahneman, 1974, 2002). Rather, it is a model which proposes that predictors will decide, consciously or semi-consciously, when they feel ready to rely on their gut feeling or to undertake more analysis. Interestingly, other studies have found that *"the degree to which people process information deliberately or intuitively has been found to depend on affective states"* (de Vries, Holland, & Witteman, 2008).

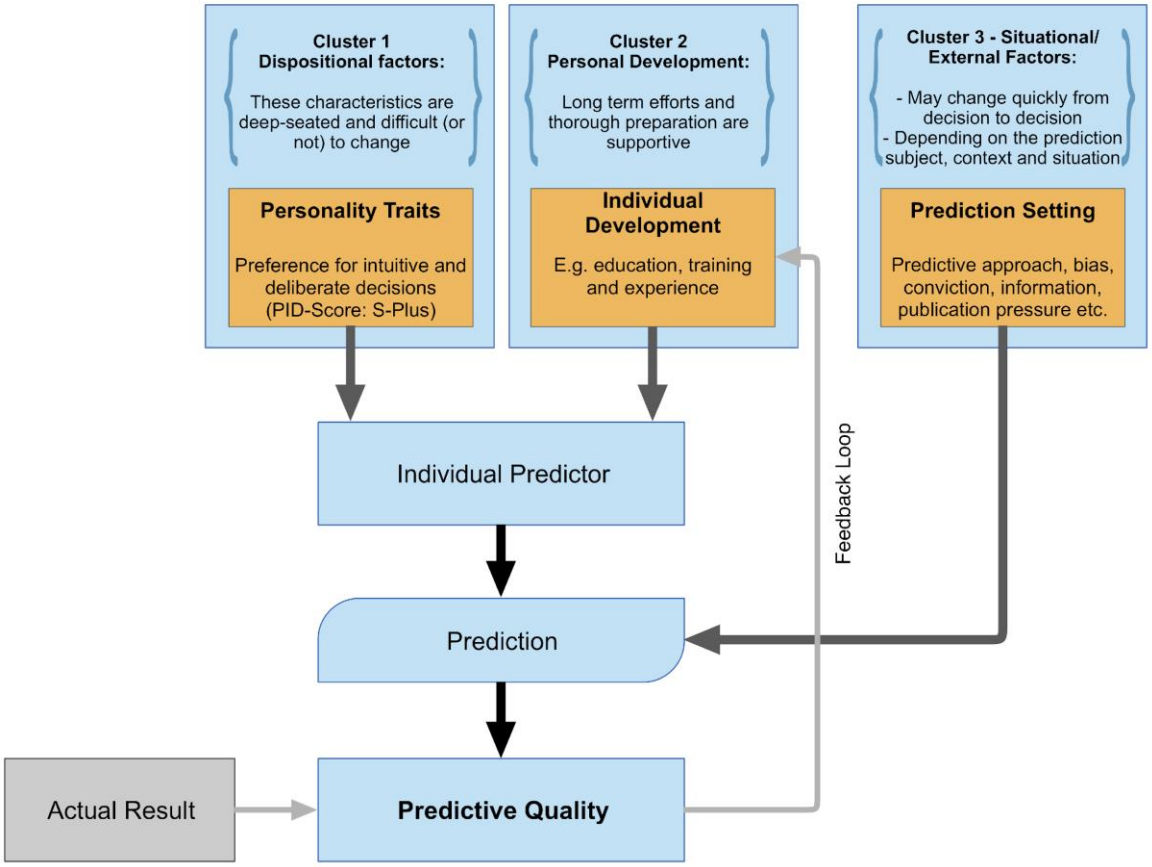

Figure 1. Deliberated Intuition Model

While no predictor was always able to provide correct predictions, there were predictors with superior performance. Additionally, the group design and group feedback loop in the experiment had only a minor impact on the individuals' predictive quality. One pattern within the group of "superior predictors" was that some individual predictor characteristics, particularly PID-S-plus score, indicated an individual preference for intuitive and deliberate decisions. However, beyond the predictor's personal predisposition, other factors contributed to good predictive quality. Participants who used an approach which could be described as informed intuitive prediction showed superior predictions. In-depth knowledge and training with deliberative forecasting methods, like formal prediction models and market experience, enhanced predictive quality. The crux of the matter is that a top predictor combines these factors with his or her personality (e.g., the PID-S-plus score), training and experience. Additionally, the setting of the actual prediction situation impacts top predictors' ways of reasoning. The deliberated intuition model can be regarded as a re-conceptualization and enhancement of dual-process models, which are sometimes able to outperform the predictive quality of financial analysts using a traditional (rational) approach. The model might help to design groups or crowdsourcing platforms but could also aid the practice and team design for traditional financial analysts.

Conclusion

Group decision-making is a complex process. In order to accommodate this complexity, a mixed-methods approach in the form of a sequential study allowed us to

address several issues. The mixed-methods approach was suitable to address the following research question: what underlying key mechanisms, for the individual and group, 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 those of 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.

Some central aspects of group decisions still seem to be the group's composition and the specific questions asked. These group decisions certainly have high value as complements to existing analysts opinions, but the results are not necessarily superior. The question of the contexts in which group decisions are particularly good, and when they are not, remains open.

Despite these open questions, online communities in the investment sector increasingly seem to be used. Several characteristics may help to enhance group design. The deliberated intuition model may help to design group settings and improve predictive quality. However, further research is needed to improve its robustness and to explore its limitations. Perhaps a practical approach is that groups can add value but only with appropriately selected individuals and settings. If used correctly, social media, particularly 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 crowds (Goldstein, McAfee, & Suri, 2014; Mannes, Soll, & Larrick, 2014). Further research should also include variations in cultural context (Wang, Rieger, & Hens, 2017), investment instruments, market conditions and assessments of additional variables such as risk aversion, trading activity, over-optimism and sensation-seeking.

References

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

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

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

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

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

Budescu, D. V., & Chen, E. (2015). Identifying Expertise to Extract the Wisdom of Crowds. *Management Science*, *61*(2), 267-280.

Camerer, C. F., Loewenstein, G., & Rabin, M. (2011). *Advances in Behavioral Economics*. Princeton University Press.

Chen, H., Hu, Y. & Huang, S. (2018). Monetary Incentive and Crowd-sourced Equity Research. *Georgia Tech Scheller College of Business Research Paper No. 17-23*, http://dx.doi.org/10.2139/ssrn.2964503.

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

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

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

Dane, E., Rockmann, K. W., & Pratt, M. G. (2012). When should I trust my gut? Linking domain expertise to intuitive decision-making effectiveness. *Organizational Behavior and Human Decision Processes, 119*(2), 187-194. https://doi.org/10.1016/j.obhdp.2012.07.009

Davis-Stober, C. P., Budescu, D. V., Broomell, S. B., & Dana, J. (2015). The Composition of Optimally Wise Crowds. *Decision Analysis, 12*(3), 130-143. https://doi.org/10.1287/deca.2015.0315

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

Dhami, S. (2016). *The Foundations of Behavioral Economic Analysis*. OUP Oxford.

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. and Gear, T. (2018) *"Deliberated Intuition for Groups": An Explanatory Model for Crowd Predictions in the Domain of Stock-Price Forecasting.* In: Proceedings of the 51st Hawaii International Conference on System Sciences 2018. Hawaii International Conference on System Sciences, Hawaii, 4094-4101. https://doi.org/10.24251/HICSS.2018.514

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

Fleischer, J. (2005). *Die Performance von Analystenempfehlungen [The performance of analyst recommendations]* (1. Aufl.). Wiesbaden: Gabler.

Gigerenzer, G. (2008). *Bauchentscheidungen: Die Intelligenz des Unbewussten und die Macht der Intuition [Gut Feelings: The Intelligence of the Unconscious]*. München: Goldmann.

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

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

Gray, L., & Diehl, P. (1992). *Research methods for business and management*. New York [etc.]: Macmillan.

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

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(46), 16385-16389.

Jackson, S. A., Kleitman, S., Stankov, L., & Howie, P. (2017). Individual Differences in Decision Making Depend on Cognitive Abilities, Monitoring and Control. *Journal of Behavioral Decision Making*, *30*(2), 209-223. https://doi.org/10.1002/bdm.1939

Jame, R., Johnston, R., Markov, S., & Wolfe M. (2016). The Value of Crowdsourced Earnings Forecasts, *Journal of Accounting Research.* 54(4), 1077-1110.

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., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

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).

Kasthurirathna, D., & Piraveenan, M. (2015). Emergence of scale-free characteristics in socio-ecological systems with bounded rationality. *Scientific Reports*, *5*, 10448.

Lin, M. (2017). The effect of 52 week highs and lows on analyst stock recommendations. *Account Finance,* https://doi.org/10.1111/acfi.12312.

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

Løhre, E., & Teigen, K. H. (2017). Probabilities Associated with Precise and Vague Forecasts. *Journal of Behavioral Decision Making*, n/a-n/a. https://doi.org/10.1002/bdm.2021

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

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

Marwala, T. (2013). Semi-bounded Rationality: A model for decision making. *ArXiv Preprint ArXiv:1305.6037*. Retrieved from https://arxiv.org/abs/1305.6037

Nguyen, M. (2018). A New Decision Making Model based on the Made in Vietnam Lean Management Philosophy. *Economics & Sociology*, *11*(1), 44-60.

Olsson, H., & Loveday, J. (2015). A Comparison of Small Crowd Selection Methods. In *CogSci*. Retrieved from https://mindmodeling.org/cogsci2015/papers/0307/paper0307.pdf

Paddock, E. L., Ko, J., Cropanzano, R., Bagger, J., El Akremi, A., Camerman, J., … Van den Bos, K. (2015). Voice and Culture: A Prospect Theory Approach. *Journal of Behavioral Decision Making*, *28*(2), 167-175. https://doi.org/10.1002/bdm.1834

Parker, A. M., & Stone, E. R. (2014). Identifying the Effects of Unjustified Confidence versus Overconfidence: Lessons Learned from Two Analytic Methods. *Journal of Behavioral Decision Making*, *27*(2), 134-145. https://doi.org/10.1002/bdm.1787

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.

Prelec, D., Seung, H. S., & McCoy, J. (2017). A solution to the single-question crowd wisdom problem. *Nature*, *541*(7638), 532-535.

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.

Prpić, J., & Shukla, P. (2016). Crowd Science: Measurements, Models, and Methods (pp. 4365-4374). IEEE. https://doi.org/10.1109/HICSS.2016.542

Prpić, J., Taeihagh, A., & Melton, J. (2015). The Fundamentals of Policy Crowdsourcing: The Fundamentals of Policy Crowdsourcing. *Policy & Internet*, *7*(3), 340-361. https://doi.org/10.1002/poi3.102

Roger, P., Roger, T., & Schatt, A. (2015). Behavioral Biases in Number Processing: The Case of Analysts' Target Prices. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2669451

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

Silver, N. (2012). *The signal and the noise: why so many predictions fail--but some don't*. New York: Penguin Press.

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

Slovic, S., & Slovic, P. (Eds.). (2015). *Numbers and nerves: information, emotion, and meaning in a world of data*. Corvallis, OR: Oregon State University Press.

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: Deutscher Universitäts-Verlag.

Steyvers, M., & Miller, B. (2015). Cognition and Collective Intelligence. *Handbook of Collective Intelligence*, 119.

Tetlock, P. E. (2005). *Expert political judgment how good is it? How can we know?* Princeton, N.J.: Princeton University Press.

Tetlock, P. E., & Gardner, D. (2015). *Superforecasting: the art and science of prediction* (First edition). New York: Crown Publishers.

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.

Wang, M., Rieger, M. O., & Hens, T. (2017). The Impact of Culture on Loss Aversion. *Journal of Behavioral Decision Making*, *30*(2), 270-281. https://doi.org/10.1002/bdm.1941. New York: Cambridge University Press.