What Drives Decision Makers to Follow or Ignore Forecasting Tools – A Game Based Analysis

Georg Vetter
TU Dortmund
georg.vetter@tu-dortmund.de

Markus Siepermann
TU Dortmund
markus.siepermann@tu-dortmund.de

Richard Lackes
TU Dortmund
richard.lackes@tu-dortmund.de

Abstract

Using forecasts is a prerequisite for good decision making but often decision makers ignore the outcomes of forecasting tools and rely solely on their personal assessment of the decision situation. On the one hand, this usually leads to worse decisions in comparison to situations where the forecast was considered. On the other hand, forecasts can also be defective. If so, decision makers are well advised not to use the forecast. Thus, it is crucial that they do not rely blindly on forecasts but scrutinize critically the results. The question is under which circumstances decision makers follow or ignore forecasts. To answer this question, we conducted a laboratory experiment where decision makers have the choice between two alternatives. The forecast provided gives an advice which alternative to choose. The forecast is manipulated so that it is only partly reliable. Results show that participants do not act optimally. If they are blinded by their success or experience several failures over time that they are not responsible for, they tend to rely on the manipulated forecast instead of calculating their own more accurate forecast.

1. Introduction

Forecasts are an inherent part of a firms's planning activities [37][51]. While planning is the process of a firm to adapt to its environment [8], forecasting is a projection into the future of an expected outcome given a stated set of environmental conditions [36]. As such, they have an important impact on the decision making process and a firm's final outcome. Usually, software tools like expert systems (ES) support the forecasting process completely or partly. However, the forecast itself is used by individuals in the decision making process so that the employment of forecasts highly depends on the involved individuals [3]. As human behavior is not only determined by objective observations and reasons but also by subjective belief of individuals [13][48], users may deviate from the objective advice of a forecast. On the one hand, users may have additional information which the forecast could not take into account so that its predictive power is limited. In this case, scrutinizing the forecast is inevitable. On the other hand, users can question the forecast if it is not in line with their expectations. There are innumerable sources for mistakes during the decision making process [10]. So if users are convinced of their abilities, they may adjust or override the forecast and act in accordance to their own beliefs [12][54].

If additional information can be excluded, the question is why individuals do not use forecasts and rely on their own estimations instead. The reasons for this seem to be manifold and to some extent contradictory. While some researchers found that the quality of a forecast significantly influences its usage [49], others could not confirm these findings [41]. The reason may lie in different fields of application, research methods, or sample data. However, a consensus is that forecasts influence the behavior of decision makers [14][23][44][52].

While forecasts are in the focus of research for many years, especially to improve accuracy, research concerning the use of forecasts and the usage reasons is scarce [2]. Behavioral sciences have emphasized the role of affect [32], emotion [29], self-confidence [6][28], self-esteem and anxiety [55] for the decision making itself and how these factors influence the choice of risky or more certain alternatives. But the role of individuals using forecasts is hardly investigated and should be put into focus [26][50]. This is particularly of interest as even decision makers who constantly and systematically make bad decisions can survive in a leadership position for a long time [11]. Therefore, this paper aims to shed light on the following research question:

RQ: Which circumstances make individuals rely on or discard forecasts in uncertain situations?

Thereby, we want to focus on the factors of the decision making itself without influences from a specific decision situation. In particular, we want to focus on the influence of experience and success. In other words, we examine how the success and the experience

of a decision maker influence his decisions regarding the usage of forecasts.

As recent investigations have shown, does the simple presence of others influence the actions of people [1][5]. If a person is surrounded by a peer group, he usually acts more risk taking than if he was alone [5]. This holds particularly if the decision maker has low self-esteem [1]. To eliminate influences of others and to focus on the role of former experience and success, we employ the concept of gamification to our study and use the simple game High or Low for our observations. As the behavior of people in games is similar to their behavior in real life [19][47], this approach provides several advantages. First of all, influences from others on the decision maker are excluded. Secondly, because we use a solo game, a human opponent also does not have any impact. And lastly, all the benefits of gamification can take effect [16][18] (see also section 2).

Within the game, the decision maker has only two options for his decision and is supported by a forecast that gives advice which alternative to choose. But the forecast is manipulated such that it gives the wrong advice in some situations. With the help of this setting we can observe if a decision maker scrutinizes the outcomes of the forecast, ignores it or follows the advice blindly. As experience is said to be important for the assessment of decision situations and of the quality of forecasts [21][49], the more experienced a decision maker is, the more he is expected to recognize the manipulation.

The remainder of this paper is structured as follows: In section 2, we give an overview about the theoretical background and related literature in this field. In section 3, we explain the methodology of our study. In particular, we introduce the game that we used to analyze the usage behavior concerning forecasts. Section 4 presents the analysis whose results are discussed in section 5. There, implications of the results are derived. In addition, some limitations and a future outlook are given.

2. Theoretical background

The paper mainly contributes to two streams of research. First of all, it analyzes the usage behavior of people concerning forecasts. For this, it secondly applies the concept of gamification to the data gathering and the execution of the study.

Games are an inherent part of our lives and exist nearly since the dawn of mankind [35]. As games are fun and usually played voluntarily and with great ambition, game concepts have been applied to many non-game applications during the past years like crowd sourcing [39], brand web sites [20], or many others

[38][46]. This process of using game design patterns in non-game contexts is usually referred to as gamification [9]. Its main purpose is to encourage users to do things, to do them more often and longer than they would have done otherwise [19][46].

In this sense, we apply the game High or Low (game concept) to the non-game context of data gathering for a research study to encourage participants to stay longer for being observed in their decision making. But if the outcome for the user is a fully-fledged game, the nature of gamification is often denied [46]. However, distinguishing a game from a non-game application is not as easy as it seems [24][25][46]. While for person A an application can be a game, for person B this may not hold. Therefore, Huotari and Hamari [24][25] used a different approach to define gamification based on service marketing:

"*Gamification refers to a process of enhancing a service with affordances for gameful experiences in order to support users' overall value creation.*" [24].

According to this definition, any service, be it a non-game application or also a game, can be enhanced by game design patterns if users experience this enhancement as an improvement and as gameful [24][25]. In this respect, the data gathering of our study is enhanced with a game for bringing a joyful experience to users while they create the data of decision making. We use the game as motivator to make the application of data gathering more interesting, more playful and more exciting for the participants so that users keep on playing/producing data.

As the use of the forecast and the decision making are the core of the game, we can focus on the risk behavior of people without having them influenced by other people. Otherwise, users would take actions to be conform with the opinion of the group [5]. Other methods could be a survey, experiments, or observations in real decision making but any of these alternative methods bears several shortcomings. In surveys, interviewees often try to comply with the views of the interviewer or other people. As the outcome is self-reported, distortions occur as people often are not able to judge their own situation or abilities correctly [27]. Experiments and observations in real life are complex and costly. Besides it is difficult to obtain a sufficient number of samples. Therefore, the game is used as a deputy for a decision situation under risk. The advantages are that the decision situation is easy [17] but always new. The motivation of the participants is kept high. The forecast can easily be manipulated so that two different situations can be realized. And lastly, although there is no group pressure, participants have the incentive to play the game seriously as they can compare their outcome to others on a leaderboard.

While other papers usually investigate how to use game elements in a non-game context in order to improve the behavioral outcome [38][46], the use of gamification for research purposes is still scarce. Besides papers in game theory who use the Prisoner's Dilemma Game and derivates to analyze risk behavior and risk strategies (e.g. [22]), we are, to the best of our knowledge, the first who employ a simple card game for their study purposes. While Mushag et al. [42] pays money incentives to participants of a lengthy survey, Rapp et al. [40] award participants of a field study with points so that they can compare their performance to the ones of other participants on a leaderboard.

Forecasts can be found in many business areas and have many different applications. Therefore, the body of literature analyzing the usage of forecasts is diverse. One stream investigates when forecasts are used by decision makers. O'Connor et al. [41] have a look at water managers and their use of climate and weather forecasts. They found that perceptions of the risk situation are influencing the usage of forecasts much more than reliability. Although managers perceive forecasts as reliable and accurate, they use them only if they are facing a risky situation or expect to face one in the near future. However, in their analysis they cannot explain more than 20% of the variance so that there must be more influencing factors that they did not examine.

Glaum et al. [15], Smith and Mentzer [49], as well as Sarens and D'Onza [45] have a look at the forecast itself. Glaum et al. [15] focus on the quality of the forecast output. They found that the effort a firm invests in the forecast, the efficiency with which the forecast is done, and the quality of the input data positively influence the outcome of forecasts. Sarens and D'Onza [45] show that when performing a forecast, analysts pay more attention to individual risks than to general risks. Smith and Mentzer [49] analyze the role that forecast accuracy plays for the usage of forecasts. They show that forecast accuracy positively influences the perceived quality and thereby the usage of forecasts by users. As they focus on logistics, they also show that the logistics performance can be improved.

Also Gaynor and Kelton [14] as well as Rupar [44] focus on the credibility of forecasts. Gaynor and Kelton [14] analyze how different forecasts of firms and analysts are perceived and used by investors. They find that if the firm's forecast is in line with the earnings trend, the analyst's forecast is perceived as less useful. Otherwise, if the firm's forecast deviates from prior trends, investors are geared to the analyst's forecast. Rupar [44] observes related results. If the forecast precision provided by firms does not meet the expectation of investors, they mistrust the forecast. This is in line with Huang [23] who found that the disclosure of

reduced forecasts dampens expectations regarding a firm's development and can limit the loss in comparison to the situation when bad news are announced.

Another stream of literature focuses on the interpretation of forecasts. Interestingly, users have problems to interpret forecasts and their own behavior correctly. Juanchich and Sirota [27] found that more than 50% of the participants of their survey are not able to interpret a forecast correctly. This is in line with Maines and Hand [34] who found that individuals do not weight time series information correctly when performing a forecast by themselves. Lucarelli et al. [33] observed that individuals could not assess their risk tolerance level correctly. Although a high share of participants of their study stated that they are risk averse, they act like a risk taker. Both phenomena may usually lead to inappropriate decision making.

Another research stream investigates characteristics of the decision maker. Lo and Repin [30] as well as Lo et al. [31] found that experience reduces emotional reactivity and improves the usage of forecasts during the decision making process.

Our paper is most related to those works that analyze the situation when the decision making takes place. In contrast to other papers, we solely consider the performance of the forecast and the experience that the decision maker made during the past periods. In addition, we control the forecast accuracy by manipulating the outcome in a certain way so that it is less reliable in some situations. This situation equals to some degree the setting of Gaynor and Kelton [14]. The forecast we provide to decision makers corresponds to the firm's forecast while the decision makers own calculations corresponds to the analyst's forecast.

3. Methodology

In order to observe the behavior of decision makers concerning the usage of forecasts, we use a laboratory experiment where participants are observed when they play the simple card game High or Low. When playing, a forecast is provided that gives advice for the next game step. This forecast is manipulated such that it gives wrong advices in certain situations. Before we explain the setting in more detail, we first have a look at the game play. For the analysis of the data retrieved, we use the data mining technique of a decision tree that is described afterwards.

3.1. The high or low game

The *High or Low* card game is one of the simplest card games played with either 32 or 52 cards. For calculation simplicity, we restrict the game to 32 cards. The order of the cards colours shall be (from high to

low) clubs, spades, hearts, diamonds, the order of the cards shall be ace, king, queen, jack, 10, 9, 8, and 7. Then, the game play is as follows: In the first step, the dealer (here: the computer player) takes the first two cards from the pile of cards and shows one card to the player. The other card is hidden. The player then gets a forecast displayed. The forecast can make three different predictions:

- Higher: The second card is probably higher than the first card.
- Lower: The second card is probably lower than the first card.
- Uncertain: The probabilities are too close to make a prediction.

Afterwards the player can choose between higher or lower. If the player is right, he gets one point for this round. If he is wrong, no points will be added to his account. At the end of this round, both cards are put on the pile with the played cards. With a deck size of 32 cards a game lasts 16 rounds. After each game, the cards are shuffled. Every player can see his own score in the game and his overall high score. All high scores are listed on a leader board.

The forecast is based on the probability that the next card is higher or lower. This probability considers all played cards and is calculated as follows:

$$probability = \#cards\ above\ /\ \#all\ remaining\ cards$$

If the ratio of the number of remaining cards above the first card divided by all remaining cards is above 0.6, the forecast advises to choose higher. On the other hand, if the ratio is under 0.4, the shown forecast is „lower". Between 0.4 and 0.6 the forecast tells that it is uncertain. The intention behind the uncertain forecast is to build trust in the forecast. Since the game is probabilistic and not deterministic, the forecast does not have to predict the correct card in each round, but in the long run an orientation to the forecast will lead to a positive game result.

To test the faith of the players in the forecast we changed its quality and inverted the probability in some cases as shown in Figure 1. Instead of providing the correct forecast, the opposite forecast was given in the range from 55%-80% and vice versa. Simply inverting the probability in every case can result in a fast detection of the manipulated forecast. Therefore, to obfuscate the manipulated forecast and to make a detection more difficult, all obvious cases were correctly predicted. To get more manipulated forecasts, we shortened the range of uncertainty. Figure 1 shows the true and the manipulated forecast probabilities.
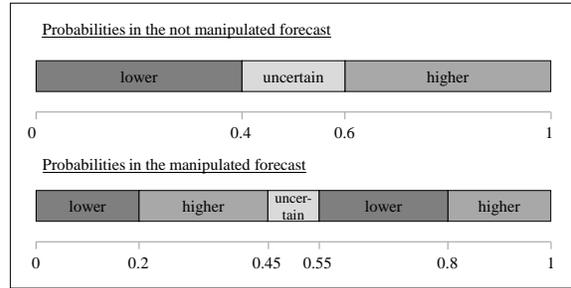

Figure 1: Probabilities of the forecast

The first not manipulated forecast scale was applied for the first two decks. The reason is that forecast accuracy plays an important role for the trust in the forecast [49]. Therefore, the first two decks act as a trust building measure so that the player gets used to an accurate forecast. From deck 3 till the end the player got just the second, manipulated forecast shown. If a player quits the game and started the game again at a later time, the first two decks again used the not manipulated forecast.

The card game was implemented as an online game playable in a web browser.

3.2. Data mining technique decision tree

The purpose of this paper is to analyze situations where players choose to follow the manipulated forecast and to derive rules as characteristics for the situations. For this, we use the data mining technique of a decision tree. A decision tree is a classifying technique that does not only classify data sets into predefined classes but also provides insights into the classifying rules.[4] Figure 2 gives an example of a decision tree.
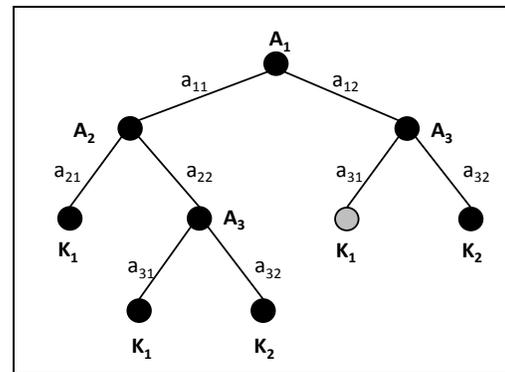

Figure 2: Example of a decision tree

There, we have two classes, K_1 and K_2. Each inner node including the root node of the tree represents one attribute of the data set. The edges represent distinct values (or distinct intervals or groups of values) of the attribute. Then, a data record is classified as follows: Starting by the root node, the dataset traverses the tree

to a leaf node. In each node, it goes down to next node along the edge that matches its own value of the attribute in the node. When the data record enters a leaf node, it is classified to the class that is indicated by the leaf node. The path from the root of the tree to a leaf represents a rule. All records classified into a leaf have taken the same path within the tree and therefore fulfil the same comparisons. In our example, the grey leaf represents the rule:

$$IF\ A_1=a_{12}\ AND\ A_3=a_{31}\ THEN\ K_1$$

The quality of a rule is indicated by two measures: confidence and support. The confidence of a rule indicates its reliability. It is calculated as the share of records classified correctly by the rule to all records classified by the rule. The support indicates how often the rule can be applied. It is calculated as the share of records classified by the rule to all records of the data set. While the confidence of a rule should be as great as possible in order to avoid faulty classifications, the support does not necessarily need to be great because due to the usually big number of data records a rule cannot be used for all situations. Instead, finding a set of reliable rules is usually sufficient if it is not necessary to classify any data record. Then, these reliable rules can be used to identify promising data records and to develop a decision strategy.

4. Analysis

We conducted the laboratory experiment in the beginning of the year 2016. Participants were mainly students who are adequate surrogates for decision makers [43] so that this sampling will hardly distort the results. In total, 112 players participated and generated 10,202 data records. Since we aim to analyze why decision makers follow or discard a manipulated forecast, we eliminated all data records where no manipulated was made. This resulted in a database of 2,626 data records.

Every record represents one playing round and contains the played cards, the hand id and the user id. To describe the success of a player his past play was classified into a success category depending on his success rate. The success rate gives no conclusion about the conformity of the decision with the forecast, but about the relationship between played rounds and won rounds.

$$success\ rate = won\ rounds\ /\ played\ rounds$$

The attribute values were defined as more successful (success rate > 0.66), medium successful (0.33 < success rate ≤ 0.66) and less successful (success rate ≤

0.33). We assumed the influence of the success category can be different according to the period under consideration. Therefore, the success category was calculated for the long-, mid- and short-term, while long-term considers the last 15 rounds, mid-term the last 10 rounds and short-term the last 5 rounds.

To describe the experience of a player, we introduced four additional attributes for every data record. The first attribute expresses the success level of the forecast, with the same values as the success category of the player and the same differentiation of periods under consideration. In addition to this, we measured the amount of played rounds by the player and the number of consecutive wins and defeats.

Table 1: Attributes

Attribute	Term	Scale
Success category user	Short-term	{low; medium; high}
	Mid-term	{low; medium; high}
	Long-term	{low; medium; high}
Success category forecast	Short-term	{low; medium; high}
	Mid-term	{low; medium; high}
	Long-term	{low; medium; high}
Number of played rounds		Continuous
Consecutive wins		Continuous
Consecutive defeats		Continuous

For classification, two classes were formed, following the manipulated forecast (Class = 1) and not following the manipulated forecast (Class = 0).

Table 2: Support and confidence

Rule number	Support	Confidence	Class
1	1%	82%	1
2	1%	82%	1
3	8%	81%	1
4	8%	80%	1
5	2%	80%	1
6	1%	80%	1
7	2%	80%	1
8	1%	76%	1
9	1%	76%	1
10	3%	75%	1
11	12%	75%	1

This data set was used to build a binary decision tree. Based on the explanation above, we focused more on a high confidence than on a high support. We set the minimum support to 1%. This means that every rule applies for at least 26 of the 2,626 datasets. In total, 38 rules were found. In our evaluation, we focused on all rules with at least a confidence of 75%.

The overall support of these 11 rules is 41%. Hence, 41% of all rounds where the manipulated fore-

cast was followed can be described by these rules with at least 75% confidence. To facilitate the evaluation, the number of played rounds was categorized into unexperienced (<100 played rounds), mid-experienced (100 to 250 played rounds), experienced (250 to 1,250 played rounds), and old stager (>1,250 rounds).

Table 3: Rule descriptions

Situation	Rule Number	Experience level	Success of User			Success of Forecast			Consecutive wins	Consecutive defeats
			Short-term	Mid-term	Long-term	Short-term	Mid-term	Long-term		
B	1	Unexperienced to experienced		High	Low to Medium			Medium	>=3.5	
A	2	Mid-experienced		Low to Medium				Medium		
B	3	Unexperienced			High			Low		0
A	4	Mid-experienced to experienced			High	Medium		Low		0
B	5	Unexperienced		High	Low to Medium			Medium		
A	6	Unexperienced to experienced			High	Medium to High	Low	Medium to High		
C	7	Unexperienced to experienced			high			Low		>=2
C	8	Unexperienced	Medium	Medium	High			Low		1
C	9	Mid-experienced to experienced	Medium	High	Low to Medium			Medium	<4	
C	10	Mid-experienced to experienced		Medium	High			Low		1
A & B	11	Mid-experienced to experienced			High	Low or High	Low to Medium	Low		0

Table 3 shows the final rules. Every attribute value is a condition of the rule. If there is an empty space, a further split of the decision tree did not increase the purity of the adjacent nodes. This means that the rule applies for all values of the attributes with the empty entry.

As we can see, the long-term success of the forecast is involved in every rule as well as the user's long-term success except of one (rule 2). By that we can assume that the long-term success has the highest influence on the decision whether to follow the manipulated forecast or not.

Also experience has an impact on the decision making. Old stagers who played more than 1,250 rounds did not fall for the manipulated forecast. Due to their experience they are aware of possibly manipulated forecasts and therefore skeptical regarding the prediction. They know the situations in which the manipulated forecast can be attractive and the outcome of following in that case. Therefore, old stage users are the only user group that is able to avoid a false forecast in any situation.

For further analysis, similar rules are aggregated. Since rules can be interpreted as situations in which the

individual has chosen to follow the manipulated forecast, the rules were interpreted and translated into situation characteristics. Due to the fact that some situations are very similar, clusters were formed which aggregate similar situations. In our evaluation three different types of situations were formed:

Situation A: The player follows the manipulated forecast because it seems that this can improve his success. Either the success of the user is permanently below the success of the forecast (rule 2) or the success of the forecast is improving (rule 4, 6 and 11). Therefore, it is comprehensible to follow the manipulated forecast.

We have to keep in mind that a manipulated forecast does not imply that the outcome is wrong. By hazard, a consecutive manipulated forecast can still have success. Rule 2 describes the situation where the mid-term success of the user is below or equal to the long-term success of the forecast. This situation can occur if the play of the user is worse than the manipulated forecasts or if the prediction just switched to the manipulated mode and the success level of the forecast is still influenced by the right predictions. In both cases it is reasonable to choose the manipulated forecast,

since it leads at least to the same success the user already has.

The argumentation for rules 4, 6 and 11 is similar to the previous one with the exception that the success of the forecast is below the success of the user. As shown in Table 3, the success of the forecast is increasing. This is enough for the user to rely on the manipulated prediction. It is very unlikely that the user can calculate the exact success rates. It is supposable that he just develops a feeling about the level of the success rates as we modeled with the three levels. If we assume that the user cannot compare close success levels, he will just notice the gain of success of the forecast. This could be an explanation for his decision. This assumption is underpinned by the observations of the following situations. Rule 11 can also be interpreted to be a characteristic of situation B.

Situation B: If the success of the player is increasing or high in the long-term, the player tends to follow the manipulated forecast. The long-term success of the player makes him careless concerning the evaluation of the forecast even if the success of the forecast is low or medium (rule 1, 3, 5 and 11).

If we look at rule 1 and 5 we can see a stagnating success level of the forecast and a rising success level of the user. The success of the user even outreached the success of the forecast. The consecutive number of wins in rule 1 underlines the high level of perceived success. Even though, the user did not question the forecast and followed the manipulated prediction. As mentioned above, it can be assumed that the user is not able to compare the exact levels of his success the one of the forecast. Therefore, he is not aware of his superior success level in comparison to the success level of the prediction. Thus, he just recognizes the increase of his own success level. A possible explanation to change the strategy at this point is the sense of security and euphoria of the user which leads to careless decisions. Due to his good performance, he takes more risks. If we take a closer look at rule 1 and 5, there is a slight difference concerning the experience of the user. According to rule 5, unexperienced users do not even need a high number of consecutive wins to fall for the manipulated prediction.

If we take a look at rule 3 and 11, we can see that users are already blinded by a win in the last round when they encounter a long-term success of their own while the success of the forecast is low in the long-run.

Situation C: Many lost rounds in the short-term force the player to change his strategy and to follow the manipulated forecast. The success of the forecast remains on a low or medium level. The setbacks in the short-term induced a nervousness which conditioned a not reflected action (rule 7, 8, 9 and 10).

The general situation is dominated by the decreasing success level of the user. The success level of the forecast stagnates at low or medium. Although the success level of the user is not falling below the success level of the prediction, the user tends to follow the manipulated forecast with his decision. This situation underpins the assumption that the user cannot compare his success level to the forecast. A possible explanation is that the user gets nervous or desperate after he recognizes his falling success level. He tries to prevent a further loss by using the manipulated forecasts. In all of four rules of situation C (rule 7, 8, 9 and 10), the number of consecutive wins and defeats does not allow a high number of wins or even demands defeats in the short-term.

5. Conclusion

5.1. Discussion

The aim of this paper was to investigate the circumstances that make individuals rely on or discard forecasts in uncertain situations. In particular, we focused on how experience and success influence false decision making. In total, we could identify eleven rules that characterize when a decision maker relies on a faulty forecast. These rules could be classified into three categories/situations.

In situation A, the decision maker either permanently performs worse than the forecast or the forecasts slightly improves over time. In situation B, the decision maker is blinded by his success in the mid- and long-term range so that he acts with less care. In situation C, the decision maker had some consecutive disappointments in the near past so that he follows the wrong forecast.

As we can see, success as well as past experiences influence the decision making process. If the decision maker performs badly, he is geared to avoiding future mistakes and relies on the wrong forecast. This result is in line with findings from behavioral sciences. If a decision maker believes in his competence, he takes more risky choices because he thinks that he can avoid losses due to his skills [28]. Vice versa, if a decision maker encounters defeats, he loses self-efficacy and self-confidence [28] so that he tends to avoid risky situations [6]. In our case, avoiding a risky situation means to follow the (wrong) forecast.

Interestingly, users do not seem to be able to remember success over a longer period. Instead, they have a diffuse impression of their success and the success of the forecast. This distorted impression interferes their ability to take the right decision.

5.2. Implications

Several lessons can be learned from this study. Expert systems are often part of the decision process in a company. Since these systems are just focusing on a small part of the real world, false forecasts are possible. Especially in a dynamic and stochastic environment forecast tools can be wrong. Problems occur in situations where the forecast gives a false prediction and the individuals which use the system solely rely on the forecast. To avoid such situations, additional information should be provided to users of a forecast. First of all, the likelihood for different possible situations should be presented. In uncertain situations, decision makers tend to rely on external advice [6]. If the uncertainty situation is described properly, users of the forecast can better assess the situation and choose the correct alternative. Secondly, ex post analyses of forecasts and the history of own decisions should be presented so that decision makers can better judge if their past decisions were correct or not and are not blinded by short-time defeats or successes.

Concerning the use of gamification in research studies, the result is promising like in many other gamification studies [46]. Although several participants did not reach the third deck and therefore did not contribute to the study, other participants used the system extensively. The group of old stagers played more than 1,250 rounds or in other words more than 78 decks. This means that with the help of gamification surveys can be made interesting for participants such that they can get into flow [7]. However, the usage of game design elements must be considered carefully. A game like we used supports the intrinsic motivation of users. If it is complemented with additional elements of external incentives like payments [40] or a lottery, these extrinsic motivational elements could hinder the success. If for example an additional lottery is used, participants may want to stop the study as soon as they fulfilled the minimum requirements for participating in the lottery.

At last, some lessons can be learned concerning experience, success, and the use of probabilities in general and in different applications. Users can easily be influenced by providing advice that seems to come from a trusted source. This holds in particular if the user encountered some disappointments in the past as he can hardly remember the whole history or success and defeat. If game designers and providers slightly manipulate game outcomes and the success of users, they can sell for example additional items in games so that the user can proceed in the game more successful. While this manipulation seems promising, it is a red flag for politician and parents. Politics should ensure that such manipulations are not legal and pursued by

law. For this, regulations are missing that ensure that such manipulations can be detected.

5.3. Limitations

As always, there are some limitations to mention. First of all, the sample size is quite small. Even though 112 users played 10,202 rounds, just 2,626 records could have been used because 70 users did not play enough rounds that they reached the third deck with the manipulated forecast.

Secondly, the study encompasses only Germany. Future research should be done in a more international context to eliminate a possible cultural bias.

Thirdly, future studies could focus on the description of the situations in more detail. The aggregation of the rules to situations is just based on observations of success and experience. By adding more dimensions to the observation, the aggregation can differ and give deeper understanding of what drives user to make false decision.

Fourthly, even if support and confidence are on a high level, the rules do not apply in each case. It is also possible that the users did not intentionally follow the forecast. False decisions in the game did not have real negative effects. Thus, the game decision situation might not be the same as the real decision situation.

Finally, one rule (no. 11) appears in two situations so that the three identified situations may not be different enough. A further examination and specification of that rule may be more appropriate.

5.4. Future work

This study will not only be continued to collect more data so that the results are based on a larger sample. In future research, we want to collect also more information of the behavior and intentions of the users. By that we want to validate our recent research and find new insights about the topic.

More dimensions should be added to the described situations to receive a deeper understanding of what drives user to follow forecasts without questioning.

It is also planned to differ the methodology by adjusting the forecast, e.g. the forecasts cannot be uncertain. Also, the player should not see the forecast automatically. Instead, he should make a request to receive a prediction. Doing so, it is possible to distinguish if the user even used the forecast and how much he relies on it.

Finally, as more data is collected, we will be able to make a user wise analysis to distinguish different player types by collecting more data. Thereby, a cluster analysis can be done to identify different types of decision makers.

References

[1] Anthony, D. B., J. V. Wood, and J. G. Holmes, "Testing sociometer theory: Self-esteem and the importance of acceptance for social decision-making", Journal of Experimental Social Psychology, 2007 (43.3), pp. 425-432.

[2] Aziz, R. A. and J.A. Manap, "Forecasting Practice and Perception of Forecasting in a Government-Funded Malaysian University", Journal of Accounting, Business & Management, 2008 (2), pp. 94-107.

[3] Berinato, S., "What went wrong at Cisco", CIO Magazine 2001, pp.52-62.

[4] Breiman, L., "Bagging Predictors" Machine learning 1996 (24), pp. 123-140.

[5] Chou, E.Y. and L.F. Nordgren, "Safety in Numbers: Why the Mere Physical Presence of Others Affects Risk-taking Behaviors", Journal of Behavioral Decision Making, 2017, (30.3), pp. 671-682.

[6] Chuang, S.-C., Y.-H. Cheng, C.-J. Chang, and Y.-T. Chiang, "The impact of self-confidence on the compromise effect", International Journal of Psychology, 2013 (48.4), pp. 660-675.

[7] Csikszentmihalyi, M., "Flow: The Psychology of Optimal Experience", Harper Perennial Modern Classics, 2008.

[8] Deshpande, R. and G. Zaltman, "Factors affecting the use of market research information: A path analysis" Journal of marketing research 1982, pp. 14-31.

[9] Deterding, S., D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining gamification", Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments, ACM, 2011, pp. 9-15.

[10] Dörner, D. and H. Schaub, "Errors in Planning and Decision-making and the Nature of Human Information Processing", Applied Psychology, 1994 (43.4), pp. 433-453.

[11] Dragota, V., "When Making Bad Decisions Becomes Habit: Modelling the Duration of Making Systematically Bad Decisions", Economic Computation and Economic Cybernetics Studies and Research, 2016 (50.1), pp. 123-140.

[12] Fildes, R. and R. Hastings, "The organization and improvement of market forecasting", Journal of the Operational Research Society 1994 (45.1), pp. 1-16.

[13] Fishbein, M. and I. Ajzen, "Predicting and changing behavior: The reasoned action approach" Taylor & Francis, 2011.

[14] Gaynor, L.M. and A.S. Kelton, "The effects of analyst forecasts and earnings trends on perceptions of management forecast credibility" Accounting & Finance 2014 (54.1), pp. 189-210.

[15] Glaum, M., K. Schnürer, and P. Schmidt, "What Determines Managers' Perceptions of Cash Flow Forecasting Quality? Evidence From a Multinational Corporation", Journal of International Financial Management & Accounting, 2016 (3), pp 298-346.

[16] Griffiths, M.D., "The educational benefits of videogames", Education and health, 2002, (20.3), pp. 47-51.

[17] Hamari, J., "Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitari-an peer-to-peer trading service", Electronic Commerce Research and Applications, 2013 (12.4), pp. 236-245.

[18] Hamari, J., J. Koivisto, and H. Sarsa, H., "Does gamification work? - A literature review of empirical studies on gamification", 47th Hawaii International Conference on System Sciences (HICSS), 2014, pp. 3025-3034.

[19] Hamari, J., K. Huotari, and J. Tolvanen, "Gamification and Economics", 2015-MIT Press Cambridge, 2015, pp. 139-161.

[20] Harwood, T. and T. Garry, "An investigation into gamification as a customer engagement experience environment", Journal of Services Marketing, 2015 (29.6/7), pp. 533-546.

[21] Higgs, B., M. Hollick, and M.J. Polonsky, "Measuring expectations: forecast vs. ideal expectations. Does it really matter?", Journal of retailing and consumer services, 2005 (1), pp. 49-64.

[22] Hogan, J. L., R. H. Risher, and B. J. Morrison, "Social Feedback and Cooperative Game Behavior", Psychological Reports, 1974 (34.3), pp. 1075-1082.

[23] Huang, W., "The use of management forecasts to dampen analysts' expectations by Chinese listed firms" International Review of Financial Analysis 2016 (45), pp. 263-272.

[24] Huotari, K. and J. Hamari, "A definition for gamification: anchoring gamification in the service marketing literature", Electronic Markets, 2017 (27.1), pp. 21-31.

[25] Huotari, K. and J. Hamari, "Defining Gamification - A Service Marketing Perspective", MindTrek '12 Proceeding of the 16th International Academic MindTrek Conference, 2012, pp. 17-22.

[26] Jones, V.D., S. Bretschneider, and W.L. Gorr, "Organizational pressures on forecast evaluation: managerial, political, and procedural influences", Journal of forecasting 1997 (16.4), pp. 241-254.

[27] Juanchich, M. and M. Sirota, "How to improve people's interpretation of probabilities of precipitation", Journal of Risk Research, 2016 (3), pp. 388-404.

[28] Krueger, N. and P. R. Dickson, "How believing in ourselves increases risk taking: Perceived self-efficacy and opportunity recognition", Decision Sciences, 1994 (25.3), pp. 385-400.

[29] Lerner, J.S., Y. Li, P. Valdesolo, and K.S. Kassam, "Emotion and decision making" Annual Review of Psychology, 2015, (66), pp. 788-823.

[30] Lo, A.W. and D.V. Repin, "The psychophysiology of real-time financial risk processing" Journal of cognitive neuroscience 2002 (14.3), pp. 323-339.

[31] Lo, A.W., D.V. Repin, and B.N. Steenbarger, "Fear and greed in financial markets: A clinical study of day-traders" No. w11243. National Bureau of Economic Research, 2005.

[32] Loewenstein, G. and J. S. Lerner, "The role of affect in decision making", Handbook of Affective Sciences, Oxford University Press, 2003, pp. 619-642.

[33] Lucarelli, C., G. Brighetti, and P. Uberti, "Misclassifications in financial risk tolerance", Journal of Risk Research, 2015 (4), pp. 467-482.

[34] Maines, L. and J.R.M. Hand, "Individuals' perceptions and misperceptions of time series properties of quarterly earnings", Accounting Review, 1996, pp. 317-336.

[35] McGonigal, J., "Reality is broken: Why games make us better and how they can change the world", Penguin Press HC, München, 2011.

[36] Mentzer, J.T., C.C. Bienstock, and K.B. Kahn, "Benchmarking sales forecasting management", Business Horizons 1999 (42.3), pp. 48-56.

[37] Moon, M.A., J.T. Mentzer, and C.D. Smith, "Conducting a sales forecasting audit" International Journal of Forecasting 2003(19.1), pp. 5-25.

[38] Morford, Z.H., B.N. Witts, K.J. Killingsworth, and M. P. Alavosius, "Gamification: the intersection between behavior analysis and game design technologies", The Behavior Analyst, 2014, (37.1), pp. 25-40.

[39] Morschheuser, B., J. Hamari, J. Koivisto, and A. Maedeche, "Gamified crowdsourcing: Conceptualization, literature review, and future agenda", International Journal of Human-Computer Studies, 2017 (106), pp. 26-43.

[40] Musthag, M., A. Raij, D. Ganesan, S. Kumar, S. Shiffman, "Exploring micro-incentive strategies for participant compensation in high-burden studies", Proceedings of the 13th International Conference on Ubiquitous Computing, ACM, 2011, pp. 435-444.

[41] O'Connor, R. E., B. Yarnal, K. Dow, K., C.L. Jocoy, and G.J. Carbone "Feeling at risk matters: water managers and the decision to use forecasts." Risk Analysis, 2005, 25(5), pp. 1265-1275.

[42] Rapp, A., A. Marcengo, L. Console, and R. Simeoni, "Playing in the wild: enhancing user engagement in field evaluation methods", Proceedings of the 16th International Academic MindTrek Conference, ACM, 2012, pp. 227-228.

[43] Remus, W. "Graduate students as surrogates for managers in experiments on business decision making" Journal of Business Research 1986 (14.1), pp. 19-25.

[44] Rupar, K. "Significance of Forecast Precision: The Importance of Investors' Expectations", Contemporary Accounting Research, 2017 (2), pp. 849-870.

[45] Sarens, G. and G. D'Onza "The perception of financial analysts on risk, risk management, and internal control disclosure: Evidence from Belgium and Italy", International Journal of Disclosure and Governance, 2017 (2), pp. 118-138.

[46] Seaborn, K. and D. I. Fels, "Gamification in theory and action: A survey", International Journal of Human-Computer Studies, 2015 (74), pp. 14-31.

[47] Sermat, V., "Is game behavior related to behavior in other interpersonal situations?", Journal of Personality and Social Psychology, 1970 (16.1), pp. 92-109.

[48] Simon, H.A., "Models of man; social and rational", 1957.

[49] Smith, C.D. and J.T. Mentzer, "User influence on the relationship between forecast accuracy, application and logistics performance", Journal of Business Logistics, 2010 (1), pp. 159-177.

[50] Stekler, H.O., "Interpreting movements in the composite index of leading indicators", Business Economics 2003 (38.3), pp. 58.

[51] Stock, J.R., and D.M. Lambert, "Strategic logistics management" 4th Ed., Boston, MA: McGraw-Hill/Irwin, 2001.

[52] Stone, D.N. "The joint effects of DSS feedback and users' expectations on decision processes and performance", Journal of Information Systems, 1995 (1), pp. 23-41.

[53] Vesa, M., J. Hamari, J. T. Harviainen, and H. Warmelink, "Computer games and organization Studies", Organization Studies, 2017 (38.2), pp. 273-284.

[54] Wheelwright, S.C. and D.G. Clarke, "Corporate Forecasting: Promise and Reality", Harvard Business Review 1976 (54.6), pp. 40-64.

[55] Wray, L. D. and E. R. Stone, "The role of self-esteem and anxiety in decision making for self versus others in relationships", Journal of Behavioral Decision Making, 2005 (18.2), pp. 125-144.