Forecasting the fast and frugal way: A study of performance and information-processing strategies of experts and non-experts when predicting the World Cup 2002 in soccer [1]

SSE/EFI Working Paper Series in Business Administration No 2003:9
May 2003

Patric Andersson
Center for Economic psychology
Stockholm School of Economics
SWEDEN

Mattias Ekman
Stockholm Health Economics
SWEDEN

Jan Edman
Penn State University
USA

Abstract

This paper investigates forecasting performance and judgmental processes of experts and non-experts in soccer. Two circumstances motivated the paper: (i) little is known about how accurately experts predict sports events, and (ii) recent research on human judgment suggests that ignorance-based decision-strategies may be reliable. About 250 participants with different levels of knowledge of soccer took part in a survey and predicted the outcome of the first round of World Cup 2002. It was found that the participating experts (i.e., sport journalists, soccer fans, and soccer coaches) were not more accurate than the non-experts. Experts overestimated their performance and were overconfident. While the experts claimed to have relied on analytical approaches and much information, participants with limited knowledge stated that their forecasts were based upon recognition and few pieces of information. The paper concludes that a recognition-based strategy seems to be appropriate to use when forecasting worldwide soccer events.

JEL Classification: C53; C90; M00

Keywords: Expert predictions; Information use; Judgmental forecasting; Overconfidence; Recognition heuristic, Sports forecasting;

---

1 An earlier version of this paper was presented at the annual meeting of Society for Judgment and Decision-making, Kansas City, November, 2002. Comments from members of the London Judgment and Decision-making Group and viewpoints from Professor Ingolf Ståhl and Dr. Elisabeth Engelberg are gratefully acknowledged.

Correspondence concerning this paper should be addressed to: Dr. Patric Andersson, Center for Economic Psychology, Stockholm School of Economics, PO. Box 6501, S-113 83 Stockholm, Sweden. Email: ppa@hhs.se. Phone: +46-8-736 9576.

1. Introduction

Experts are often consulted about the probable course of future events in various areas. Scientific evidence suggests, however, that forecasts based on expert judgment are seldom accurate. In his book on future-predicting experts, Sherden (1998) concluded that meteorologists were not always correct, but had by far the best accuracy compared to economists, stock-market analysts, population researchers, management prophets, and social-trend spotters. Research shows that experts predict stock-markets (e.g., De Bondt, 1991) as well as business trends (e.g., Mills & Pepper, 1999) poorly. Another area, where expert judgment plays a major role, is forecasting outcomes of sports events. This area has not been extensively investigated (cf. Boulier & Stekler, 1999). The few studies of sports experts have given the following findings: (1) newspaper tipsters are more successful in predicting soccer matches than the simple strategy of assuming home wins (Forrest & Simmons, 2000); (2) expert gamblers are able to pick more winning horses than random selection (Ladouceur, Giroux & Jacques, 1998); and (3) predictions of football editors tend to be outperformed by those of rankings, a simple model, and the betting market (Boulier & Stekler, In press). Rankings (seedings) have been shown to be good predictors of matches in basketball, tennis, and football (Boulier & Stekler, 1999, In press).

An additional circumstance motivating the present paper is a recent trend in the study of human judgment. Since the 1970s, research on human judgment has been largely dominated by one research program, which focuses on investigating the fallibility of using cognitive shortcuts or heuristics in judgment and decision-making (cf. Kahneman, Slovic & Tversky, 1982). A relatively new strand of research, which is gaining increased popularity among researchers in various fields, takes the opposite approach by assuming that heuristics often lead to adequate judgments. This strand is referred to as the research program on fast and frugal heuristics (Gigerenzer, Todd & the ABC group, 1999). Among other things, this program argues that by employing simple decision-strategies individuals with limited knowledge are able to arrive at equally, or even more, correct predictions than individuals who have extensive knowledge. In other words, experts could be beaten by non-experts. Little is known about whether such simple decision-strategies apply to judgmental forecasting. The present paper attempts to give some contributions to this issue.

In the present paper we report on an empirical study of how people with different levels of knowledge predicted the outcome of the first round of the World Cup 2002 in soccer. This worldwide sports event took place in June 2002 in Japan and Korea. The paper aims to provide insights into the forecasting performance and judgmental processes of soccer experts and non-experts. In particular, we investigate the following research questions: How good are experts at forecasting soccer? Are their forecasts better than those made by non-experts? How do experts and non-experts process information when forecasting soccer? Are there any information-processing strategies that successfully simplify forecasting soccer?

The remainder of the paper is organized in four parts. First, the research on expert judgment and the research on heuristics are briefly reviewed. Second, the method is described and details are given about the participants. Third, the results are reported. Fourth, the paper ends with conclusions.

2. Review of Literature and Hypotheses

2.1. Research on expert judgment

The role of expert judgment has been extensively investigated. A very large body of research suggests that experts in various domains seldom generate better predictions than non-experts who have received some training, and that the predictions of experts are completely outperformed by those made by simple statistical models (Camerer & Johnson, 1991). For instance, in the area of predicting corporate earnings, a model, which simply assumes that there will be no changes, may result in more accurate forecasts than expert judgment and advanced statistical models (Conroy & Harris, 1987). Random-based benchmarks have also been used to evaluate expert judgment. For instance, the stock-picking ability of financial experts has been pitted against stocks picked by darts thrown at the stock tables; such a comparison indicated that financial experts performed significantly better than chance (Sundali & Atkins, 1994). It follows that good performance of experts can mean a combination of three possibilities: (i) experts are better than chance, (ii) experts are better than non-experts, or (iii) experts are better than a simple model. In a sense, the present paper attempts to find the combination that is appropriate to use when it comes to experts in soccer.

A commonly employed approach to assess the quality of expert judgment is calibration (cf. Bolger & Wright, 1992). In essence, this notion assumes that the confidence an expert puts in his/her judgments concerning a certain domain should reflect his/her ability to form correct judgments in that domain. A match between stated confidence and actual ability means that the expert is well calibrated and aware of the limits of his/her knowledge (cf. Russo & Schoemaker, 1992). An overwhelming body of research shows that most experts overestimate their ability to perform accurately, meaning that experts are overconfident (e.g., Ayton, 1992; Bolger & Wright, 1992; Allwood & Granhag, 1999). Some exceptions exist: weather forecasters and expert bridge players appear to be well calibrated (Bolger & Wright, 1992).

Thus, the notion of overconfidence has played an important role in research on expert judgment. Providing an extensive description of this notion is beyond the purpose of the present paper, but it should be noted that access to (additional) information makes people more confident (Oskamp, 1982). As noted by Ayton (1998) overconfidence has mainly been studied using general knowledge questions (e.g., which country, Indonesia or Sudan, has a population with a higher mean life expectancy?). It is reasonable to expect that this phenomenon exists in the context of judgmental forecasting. For instance, studies have shown that managers exaggerate the accuracy of their forecasts (Aukutsionek & Belianin, 2001) and that the behavior of online investors is compatible with overconfidence (Barber & Odean, 2002).

2.2. Research on fast and frugal heuristics

Recently, research on human judgment has taken another perspective. Instead of studying how human judgment leads to inferences that systematically deviate from rational models of economics (cf. Kahneman, Slovic & Tversky, 1982), a new research program has focused on investigating the efficiency of using cognitive shortcuts when making judgments and decisions (Gigerenzer et al., 1999). This program emphasizes that heuristics are efficient, because they enable people to quickly arrive at (frequently) accurate conclusions despite constraints in information-processing capacities, limited knowledge, and ambiguous environments. Heuristics should be evaluated with respect to their fit with the structure of task environment rather than with the rational model of economics (Gigerenzer et al., 1999). In principle, the structure of task environment concerns the relation between a criterion value (e.g., corporate earnings) and its predictors (e.g., characteristics of board members, economic growth) as well as the interrelations of the predictors. For instance, in a task environment characterized by non-compensatory information, where one cue may be more important than any combinations of other less valid cues, a judgmental strategy that considers only one cue leads to surprisingly accurate predictions (Martignon & Hoffrage, 1999). Such a strategy is deemed to be ecologically rational, as it successfully exploits the characteristics of the task environment (Gigerenzer et al., 1999).

Moreover, the program claims that people rely on a set of fast and frugal information-processing strategies when making judgments and decisions. The strategies are fast and frugal in that they demand a minimum of time, knowledge, and cognitive activities (Gigerenzer et al., 1999). The set is referred to as the "adaptive toolbox", because the use of the different strategies depends on individual factors (e.g., experience) and environmental characteristics (e.g., ambiguous information). The toolbox involves three types of fast and frugal strategies (Gigerenzer & Goldstein, 1996).

First, *the recognition heuristic* is the simplest as well as fastest of them and takes advantage of ignorance or limited knowledge. To employ the recognition heuristic, it is required to have limited knowledge of a domain (e.g., soccer). Once an individual becomes familiar with a domain his/her judgment will no longer be based upon mere recognition, since he/she will then consider additional cues of information (Goldstein & Gigerenzer, 2002). Research also suggests that more experienced individuals tend to search for and consider a greater number of cues than inexperienced individuals (Andersson, In press). The recognition heuristic works as follows: "If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion" (Goldstein & Gigerenzer, 2002, p. 76).

Some empirical evidence suggests that recognized-based judgment can be accurate. An unpublished study, which relates to the present paper, reported that individuals with limited knowledge performed equally well as knowledgeable individuals in forecasting the outcomes of (English FA-Cup) soccer matches (Ayton & Önkal, 1997). The explanation was that the former group of individuals used recognition as a cue for soccer team performance. (The knowledgeable group searched for other cues.) The recognition heuristic seems also to be reliable in the domain of picking stocks (Borges, Goldstein, Ortmann & Gigerenzer, 1999) One should note that in these two cases there has been a strong correlation between the recognized objects (i.e., teams and stocks) and the criterion (i.e., good performance). If such a strong correlation had not existed, the individuals with limited knowledge would probably have done worse. Thus, the recognition heuristic is not universally applicable (Goldstein & Gigerenzer, 2002).

Recognition relates to the notions of availability and familiarity, but there are differences. Availability concerns searches in memory for events that come easy to mind or are salient. For instance, when assessing the likelihood of business success, one might recall instances of entrepreneurs who have been successful (or have failed) and, accordingly, rely on the well-know availability heuristic (cf. Tversky & Kahneman, 1974). Familiarity assumes a continuum of knowledge (e.g., years of experience), while recognition claims only two levels of knowledge; either an object is recognized, or it is not, and there is no need for knowledge beyond recognition (Goldstein & Gigerenzer, 2002).

The two other types of fast and frugal strategies are similar and called *take-the-best heuristic* and *the minimalist heuristic.* The former type means that the decision-maker deliberately searches for the cue that perceivably best discriminates between an array of recognized objects or events (Gigerenzer et al., 1999). Once the best cue has been identified, the decision-maker is assumed to employ it when dealing with that particular task. The minimalist heuristic assumes that the decision-maker considers the information about the recognized objects or event in a non-systematic order and attempts to find the first cue that seems to differentiate between the objects or the events (Gigerenzer et al., 1999). Note that the difference between the take-the-best and the minimalist heuristics is that the former aims at finding the cue that appears to best separate an array of recognized objects or events. Czerlinski, Gigerenzer, and Goldstein (1999) have found that both types of heuristics generally perform better (i.e., yield more correct classifications) than multiple regression models involving several cues. The reason for this surprising result is that these heuristics can successfully exploit the characteristics of information.

2.3. Hypotheses

On the basis of the reviewed literature, the following hypotheses were formulated:

(1) Experts and non-experts will forecast the World Cup in soccer equally well.

(2) Experts will tend to be overconfident as regards their forecasting ability.

(3) Experts and non-experts will be outperformed by a simple model.

(4) Individuals with limited knowledge will rely on the recognition heuristic.

Moreover, we will analyze and describe other psychological aspects of forecasting such as confidence, judgmental processes, and information use.

3. Method

The paper is based on data collected by questionnaires in May 2002 and before the start of the World Cup in soccer. Participants forecasted the outcome of the first round of this worldwide sports event involving 32 national teams. Thus, the task was to individually predict which 16 teams that would play in the second round.

3.1. Questionnaire

There were two types of questionnaire. One type (Form A) had no information, while the other type (Form B) provided seven cues about the 32 qualified national teams. The cues were: number of soccer players and soccer clubs in the different countries, odds, previous participation, competition records, world ranking, and qualification records. This experimental manipulation made it possible to investigate how access to information affected performance, confidence, and judgmental processes (see below for more details). The two forms, which consisted of seven or nine pages, proceeded as follows.

- Questions with regard to how well the participant knew the 32 national teams qualified to the World Cup. The teams were presented in random order. Tests showed that there was no order effect. Knowledge of each team was rated on a seven-point verbally anchored scale ranging from “not at all” (0) to “very well” (6). The Cronbach’s alpha coefficient was 0.98 indicating extremely good reliability.
- Questions about predictions of the outcome of the first round. The participant was asked to predict the two teams in each of the eight playing groups that would qualify to the second round. (Each playing group had four teams). If the participant had Form B, he/she obtained cues about the teams. For each prediction, confidence in the predicted teams was measured on a seven-point verbally anchored scale ranging from “very uncertain” (0) to “very certain” (6). Accordingly, there were eight confidence scales.
- A question concerning the notion of confidence. Specifically, the participant was asked to state how many of the 16 selected teams he/she believed would actually qualify.[2]
- A question about the judgmental process. The participant rated how well he/she agreed with seven statements concerning different information-processing strategies like the three

[2] After this question, the forms continued with two minor sections. The first section concerned forecasts of the two teams playing the World Cup final and the world champion, respectively. Confidence in these forecasts was rated on a seven-point scale. The second section involved two questions of how important it was for the participant to be correct and to what extent the predictions agreed with other participants. The responses of these sections were disregarded as they were deemed to be irrelevant for the purpose of the present paper.

heuristics in "the adaptive toolbox" (cf. Gigerenzer et al., 1999) as well as extensive use of information. Agreement with each of the seven statements was evaluated on a seven-point verbally anchored scale ranging between "no, absolutely not" (-3) and "yes, absolutely" (3). The scale's midpoint (0) was denoted "doubtful".

- Questions with regard to use of information. The participant reported to what extent 21 pre-specified cues (presented in random order) were considered when he/she predicted the outcome of first round. Tests showed that there was no order effect. The use of each cue was rated on a seven-point verbally anchored scale with the endpoints "not at all" (0) and "to a very large extent" (6). The Cronbach's alpha coefficient was 0.89 indicating very good reliability. The participant could also add other cues.
- Questions about interest in soccer, knowledge of soccer, ability to forecasting soccer, and related issues (e.g., number of followed World Cups and number of watched soccer matches on-site during the last twelve months). Interest, knowledge, and forecasting ability (in comparison to soccer-interested people) were rated on three separate seven-point scales with different anchors which ranged as follows: "completely uninterested" vs. "extremely interested", "very poor" vs. "very good", and "much worse vs. "much better", respectively.
- Questions with regard to demographics, motivation to participate, (self-estimated) response time, and, when appropriate (see below), choice of lottery prize (i.e., incentive). About 70% of the participants reported that they were motivated. On average, it took roughly 12.7 minutes (SD = 6.6) to answer the questionnaire.

3.2. Participants

Form A was sent out to various Swedish soccer experts, while both types of questionnaire were handed out to Swedish students and American students. In all, there were 251 participants.

Experts: This group involved 52 (all male) participants with expertise in soccer. These participants belonged to three subgroups: sports journalists, soccer fans, and soccer coaches. Apart from diverse biographical backgrounds, these subgroups gave similar responses. Thus, there were essentially no significant differences. To facilitate the analyses, the subgroups were amalgamated.

About 40 sports journalists and commentators specialized to cover soccer in the major Swedish newspapers and the Swedish television, were contacted by phone and mail and asked to participate. Given that the questionnaire was sent out just three weeks before the World Cup started, the response rate of 48% must be viewed as high. Thus, 19 (all male) journalists and commentators, henceforth referred to as sports journalists, returned completely responded forms. Their mean age and mean experience of sports journalism were 43.8 and 18.3 years. On average, they had previously covered three World Cups in soccer.

The second subgroup included soccer fans, which were recruited by two approaches. First, questionnaires were distributed among supporters of one soccer club playing in the Swedish premier league. Second, a sort of snowballing procedure was used, meaning that questionnaires were sent out to various people whom were known by the authors to have a deep interest in soccer and possess great knowledge of soccer. The selected people were also asked to give questionnaires to individuals whom they believed were equally true soccer fans. These two approaches resulted in 40 respondents. To make sure that these respondents could actually be classified as true fans, control was made with respect to their self-reported interest in soccer. Those respondents who stated that they were very or extremely interested were categorized as true soccer fans. In all, 22 (all male) respondents were identified as being true soccer fans and their mean age was 33.8 years. The responses of the remaining participants were disregarded.

The third subgroup of experts involved soccer coaches. About 40 coaches and managers of clubs in the Swedish premier league were contacted by mail and asked to respond to a questionnaire. Probably due to the fact that the questionnaire was sent out only nine days before the kick-off of the World Cup, the response rate was modest. Eleven (all males) coaches and managers responded. On average, they were 44.6 years old and had 16.1 years of experience in coaching soccer clubs.

Swedish students: In all, 167 (106 males) students from Stockholm School of Economics (SSE) participated on a purely voluntarily basis. Their average age was 23. Because the majority of the students at SSE are men, it was not surprising that there were more males among the participating students, who were randomly given the two types of questionnaire. The resulting two groups were then divided with respect to the responses to the aforementioned question about knowledge of soccer. This procedure resulted in the following four subgroups: (i) 30 (25 male) knowledgeable students with information, and (ii) 44 (36 male) knowledgeable students with no information, (iii) 38 (19 male) naïve students with information, and (iv) 54 (26 male) naïve students with no information. Two students did not return complete forms.

American students: Forty-one (26 males) students from Penn State University took part on a purely voluntarily basis. Their mean age was 23.5. The two types of questionnaire were randomly distributed among them. Eight students were excluded, as they did not return completely answered forms. As a result, the two groups were as follows: (i) 18 (12 males) students without information and (ii) 15 (10 males) students with information.

3.3. Procedure

Except for the sports journalists and the soccer coaches, all participants were offered incentives for their participation and were told that they could win prizes in two ways: by making the most accurate forecasts and by just taking part. The purpose of those prizes was to trigger the participants to make an effort to perform well. For example, the Swedish participants had the choice of two prizes (1 cinema ticket or 3 lottery tickets) worth approximately USD 8 (55% chose the cinema ticket). The sports journalists and the soccer coaches were not offered such prizes, because as experts they were assumed to be driven by intrinsic motivation to participate and predict accurately without additional incentives (cf. Shanteau & Stewart, 1992).

As mentioned above, the participants were categorized into four groups: experts, knowledgeable Swedish students, naïve Swedish students, and American students. These groups were argued to correspond to four levels of knowledge ranging between expertise and (almost) ignorance. In contrast to the experts, the three other groups of participants were exposed to an experimental manipulation: absence vs. presence of information. Apart from the effects on performance and confidence, this treatment variable was assumed to facilitate the use of various information-processing strategies. For instance, participants with limited knowledge might rely on the recognition heuristic when faced with no information, whereas they might be tempted to use another strategy when faced with information. Thus, there were two independent variables: level of knowledge and access to information.

4. Results

4.1. Reported knowledge of the teams

Before making the forecasts and, when applicable, before being exposed to the experimental manipulation, the participants reported how well they knew each of the 32 teams in the World Cup. These variables were aggregated into a measure called team knowledge. In order, the mean values of this measure for the experts, the knowledgeable Swedish students, the naïve Swedish students, and the American students were 3.42, 2.84, 1.22, and 0.53 (SDs = 0.13, 0.11, 0.10, and 0.16). As expected, the four groups differed significantly with respect to team knowledge ($MS = 93.82$, $F(3, 250) = 105.89$, $p < 0.001$). Tamhane's post-hoc test showed the following significant differences: (i) the experts knew the teams greater than the knowledgeable Swedish students (mean difference = 0.60, $p < 0.01$), the naïve Swedish students (2.20, $p < 0.001$), and American students (2.90, $p < 0.001$); (ii) the knowledgeable Swedish students had greater team knowledge than the naive Swedish students (1.60, $p < 0.01$) and American students (2.30, $p < 0.001$); and (iii) the naïve Swedish students knew the teams better than the American students (0.69, $p < 0.01$). Further details about how well the participants knew the teams are given in Appendix A.

4.2. Forecasting performance and confidence of the participants

As shown by the first column in Table 1, the participants who knew less were not outperformed by those who knew more, implying support for the first hypothesis. Despite their limited knowledge, the American students and the naïve Swedish students managed to make slightly, but not significantly, better forecasts than the experts; median number of correct forecasts was 10 and 9, respectively.[3] Participants who obtained information about the teams did not outperform those who had no such information.

As can be seen in Appendix B, the participants made similar forecasts despite having different levels of knowledge and access to information. For instance, an overwhelming majority predicted that France, Spain, and Brazil would qualify to the second round, while few believed that Senegal and Saudi-Arabia would survive the first round.

[3] One might argue that the poor performance of experts was due to upsets, where three big favorites were beaten in the first round. Even when consideration was taken to these upsets, the experts did not outperform the participants with limited knowledge.

Comparing the first and the second columns in Table 1 results in a measure of calibration, which is an approach to assess the quality of forecasting. If the participants were well-calibrated and had insights into their forecasting ability, judged accuracy would be related to actual number of correct forecasts. As reported in the third column in Table 1, Wilcoxon signed-rank tests showed that the two variables were significantly different concerning all categories except for the naïve Swedish students. This finding indicates that the experts and the knowledgeable Swedish students overestimated their ability to produce accurate forecasts, whereas the American students seemed to put little faith in their predictions.

Table 1. *Forecasting performance and confidence of participants.*

	Actual number of correct forecasts. Median	**Judged number of correct forecasts. Median**	**Calibration. Mean (SD)**	**Aggregated confidence in the predictions[c] Mean (SD)**	**Percentage of participants with fore-casting ability above vs. below average[d]**
Experts (n = 52)	9	11	2.16 b (1.66)	3.88 (0.80)	54% vs. 2%
Knowledgeable Swedish students with information (n = 30)	9.5	12 a	1.89 b (1.95)	4.35 (0.82)	47% vs. 17%
Knowledgeable Swedish students with no information (n = 44)	9	12 a	2.68 b (2.23)	4.04 (0.92)	36% vs. 25%
Naïve Swedish students with information (n = 38)	10	10	0.37 (3.14)	3.13 (1.25)	5% vs. 89%
Naïve Swedish students with no information (n = 54)	10	10 a	-0.83 (3.12)	2.47 (1.51)	2% vs. 94%
American students with information (n = 15)	10	8 a	-1.53 (3.81)	3.33 (1.11)	7% vs. 79%
American students with no information (n = 18)	10	6 a	-3.56 b (2.71)	1.90 (1.67)	11% vs. 78%

Note. It was possible to have 16 correct forecasts, which no participants managed to have.
Calibration was the difference between judged and actual number of correct forecasts.
a = Wilcoxon signed-rank test indicated strongly significant ($p < 0.001$) differences between actual and judged number of correct forecasts.
b = One sample T-tests showed that these values differed significantly from zero ($p < 0.001$).
c = The measure was calculated by averaging the eight confidence scales across the eight World Cup groups. The range was between 0 (“very uncertain”) and 6 (“very certain”). The midpoint (3) meant “doubtful”.
d = The discrepancy corresponded to the proportion of participants reporting that their forecasting ability was neither above nor below average.

Recall that confidence in the forecasts was stated on eight separate scales. Aggregating the responses to these scales resulted in a measure called aggregated confidence. As shown by the fourth column in Table 1, the experts and the knowledgeable Swedish students tended to have high scores, while the students with limited knowledge had low scores. Another measure of confidence was the participants' own rating of their forecasting ability. The fifth column shows that 54% of the experts thought that they were better at predicting than their colleagues. In short, Table 1 suggests that those who knew most tended to be overconfident, while those who knew least seemed to be underconfident. Thus, the second hypothesis was supported.

In conjunction, it should be noted that the variable representing team knowledge was positively correlated ($r_s = 0.49$) with judged number of correct forecasts but negatively correlated ($r_s = -0.15$) with number of correct forecasts. There were positive correlations between this variable and the measures of calibration and aggregated confidence ($r_s = 0.49$ and 0.53).

The effects of access to information and knowledge were tested using two ANOVA models, which involved two between-subjects factors and the interaction of those factors. The first independent variable concerned absence or presence of information. Reflecting three levels of knowledge, the second independent variable consisted of the knowledgeable Swedish students, the naïve Swedish students, and the American students. Recall that the experts were not subject to the experimental manipulation, so they were excluded in the following analyses. The first ANOVA-model, which had calibration as the dependent measure, resulted in a strongly significant between-subjects factor with respect to knowledge ($MS = 27.35$, $F(3, 186) = 33.73$, $p < 0.001$), a non-significant between-subjects factor concerning information, and a significant interaction ($MS = 2.78$, $F(2, 186) = 3.43$, $p < 0.05$).[4] Tamhane's post-hoc test showed that: (i) the knowledgeable students differed from the naïve Swedish students (mean difference = 0.86, $p < 0.001$) and the American students (1.59, $p < 0.001$); and (ii) the naïve Swedish students were different from the American students (0.73, $p < 0.01$).

In the second ANOVA-model with the aggregated confidence as the dependent measure, there were two significant between-subject factors regarding information (MS =

[4] The problem with unequal observations was managed by randomly dividing the sample so that the factors consisted of similar numbers of observations. To test the robustness of the results, this procedure was repeated several times. The reported statistics concern the models run on all observations. The models based on randomly divided samples had similar statistics.

13.75, F (1, 198) = 16.73, p < 0.001) and knowledge (MS = 25.29, F (2, 198) = 30.77, p < 0.001), but a non-significant interaction. The Tamhane's post-hoc test showed that the knowledgeable Swedish students differed from the American students (1.42, p < 0.001). Thus, both knowledge and access to information lead to increased levels of confidence.

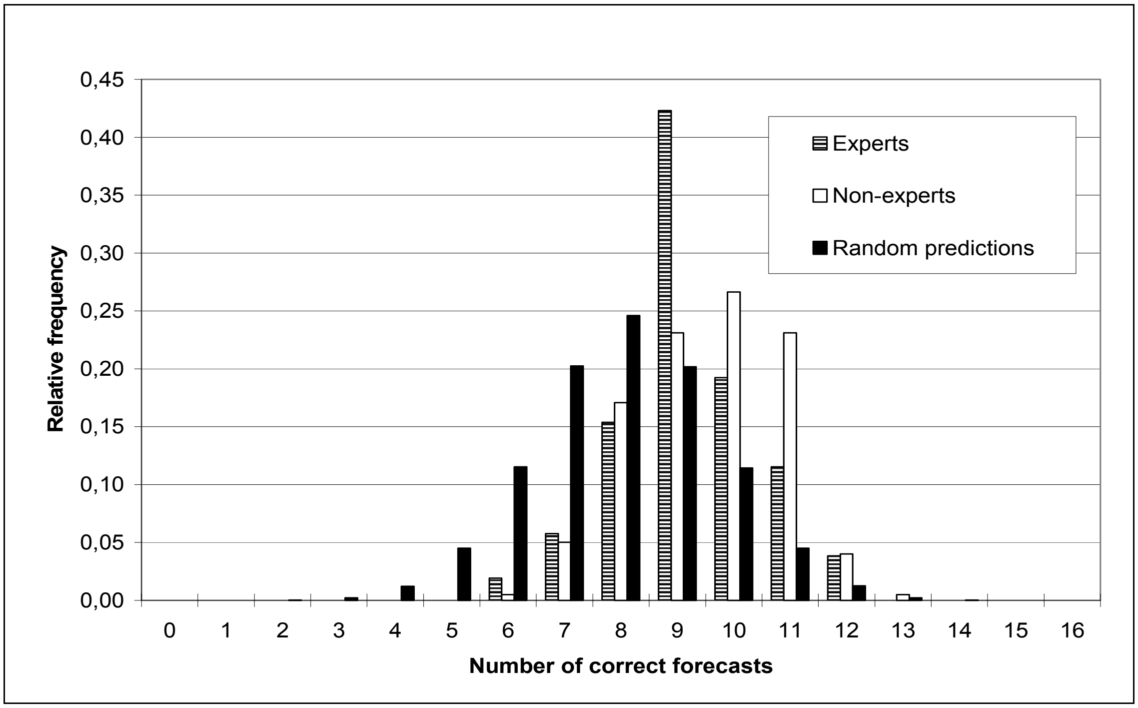

Figure 1. Relative frequencies of distribution concerning (actual) numbers of correct forecasts made by the experts, the non-experts, and the random-based model, which was constructed by simulating predictions.

4.3. Comparing forecasting performance of participants and models

The performance of the participants was also pitted against three different models working as benchmarks. First, a random-based model was constructed by simulating 500,000 predictions. Figure 1 describes the results of these simulations as well as the actual number of correct forecasts by the experts and the non-experts. The probability to have nine or more correct forecasts was 0.38 given random guessing. The majority of the participants had such numbers of correct forecasts. Note that the proportion of non-experts with more than nine correct forecasts was larger than the proportion of experts with the corresponding numbers. About 23% of the experts and 23% of the non-experts did not perform better than the random-based benchmark. On the whole, Figure 1 indicates that the seven categories of the participants performed better than chance.

Second, a simple model based on world rankings was built. Specifically, this model assumed that the two teams with the highest world rankings were selected in the playing groups to qualifying to the second round. The model was reliable and resulted in 12 correctly judged teams, which was significantly more than what most participants had.

Table 2. *Knowledge of the teams in the World Cup as reported by the nine categories of participants.*

	Knowledge of predicted		**Knowledge of actual**	
	Winning teams Mean (SD)	**Losing Teams Mean (SD)**	**Winning teams Mean (SD)**	**Losing Teams Mean (SD)**
Experts	4.19 (1.22)	2.64 *** (1.38)	3.87 (1.21)	2.97 (1.39)
Knowledgeable Swedish students with information	3.70 (1.23)	2.00 *** (1.16)	3.27 (1.48)	2.42 (1.36)
Knowledgeable Swedish students with no information	3.61 (1.26)	2.01 *** (1.26)	3.33 (1.49)	2.28 * (1.31)
Naïve Swedish students with information	1.55 (0.95)	0.82 ** (1.02)	1.56 (1.21)	0.81 * (0.68)
Naïve Swedish students with no information	1.72 (0.90)	0.78 *** (0.87)	1.59 (1.12)	0.90 * (0.74)
American students with information	0.31 (0.14)	0.19 ** (0.06)	0.27 (0.15)	0.23 (0.10)
American students with no information	0.98 (0.38)	0.48 *** (0.20)	0.88 (0.47)	0.64 * (0.28)

Note. The measures have been calculated by taking the average values representing knowledge across the teams. The range was between 0 (no knowledge) to 6 (great knowledge). Mann-Whitney U-tests analyzed the differences between winners and losers.
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Third, the team knowledge of the participants was used. Compared to the teams judged to be losers, the teams predicted to qualify to the second round (i.e., judged winners) were associated with greater levels of knowledge among the participants (see Table 2). Similarly, the participants appeared to know the actual winners better than the actual losers. Apparently, the (actual and expected) achievement of a national team was related to how well the participants knew the team. To what extent this tendency could improve forecasting was investigated by comparing models that made predictions on the basis of how well the participants knew the different teams. One such model was able to correctly classify 12 out of the 16 teams that went to the second round. This model assumed that the successful teams were the ones of which the naïve Swedish students knew best in each playing group. Models that followed the team knowledge of the American students or the experts yielded 9 and 11 correct forecasts.

In conclusion, the simple model of rankings and the recognition-based model yielded more correct predictions than the experts and the non-experts. Thus, the third hypothesis was supported.

Given the results in the Table 2, one might speculate that the participants made their forecasts with respect to their knowledge of the teams in the sense that they reckoned that the teams they knew best would qualify to the second round. This speculation is tested below.

Table 3. *Mean ratings of the seven statements concerning information-processing strategies.*

	Statement 1	**Statement 2**	**Statement 3**	**Statement 4**	**Statement 5**	**Statement 6**	**Statement 7**
Experts	1.02	-0.49	-0.58	-0.86	-2.06	0.92	1.08
Knowledgeable Swedish students with information	-0.57	-0.40	-0.80	-0.97	-2.13	0.69	0.24
Knowledgeable Swedish students with no information	1.45	0.00	-0.36	-0.64	-2.23	0.68	1.45
Naïve Swedish students with information	-0.42	1.26	0.08	-0.11	-1.58	0.37	-0.21
Naïve Swedish students with no information	0.00	1.63	-0.70	-1.09	-1.11	-0.04	-0.26
American students with information	1.13	0.33	0.20	0.73	-0.73	0.93	0.87

Note. Data were not collected for the American students obtaining no information, because it was deemed that they would have difficulties in answering questions about their judgmental processes. The participants rated how well they agreed with the statements on a bi-polar seven-point scale ranging from -3 to 3. A value above the scale's midpoint of 0 indicates agreement with the statement. The precise wording of the seven statements was as follows:

1 = I tried to analyze all available items of information concerning the nations or the teams and weigh those items.
2 = I chose those nations or the teams I had heard of earlier with regard to soccer.
3 = I tried quickly to find information about the differences between the nations' or the teams' qualities.
4 = I tried to find the item of information which best describes the differences between the nations' or the teams' qualities.
5 = I chose the nations or the teams arbitrarily.
6 = I relied on the item of information I believe is important in these circumstances.
7 = I relied on all conceivable items of information I believe are important in these circumstances.

The seven statements had varying levels of inter-correlation and could be condensed into three factors. One factor included the first, sixth, and seventh statements. The second factor incorporated the third and fourth statements. The third factor consisted of the remaining statements.

4.4. Judgmental processes of the participants

The participants were asked to evaluate seven statements representing different approaches to making forecasts and processing information (see Table 3). Regardless of whether or not information was present, the naïve Swedish students stated that they had employed a kind of recognition heuristic when making the forecasts. The experts, the knowledgeable Swedish students with no information, and the American students with information claimed that they relied on extensive analyses of information. Consequently, the fourth hypothesis was partly supported. On average, there was no agreement with the statements that represented the fast and frugal heuristics called minimalist (statement 3) and take-the-best (statement 4). Few participants reported that their forecasts were based upon random selection of teams. The seven statements were uncorrelated with the number of correct forecasts

Table 4. *Rotated factor pattern of the ratings of the information items.*

Variable	**Factor 1 (21%)**	**Factor 2 (18%)**	**Factor 3 (13%)**	**Factor 4 (7%)**
7. Star players and celebrities in the different national teams	*0.82*	0.01	0.31	-0.02
8. The quality of the different nations' soccer leagues	*0.77*	0.10	0.13	0.03
9. Previous participation in Soccer World Cups *	*0.76*	0.07	0.21	0.09
2. The soccer tradition in the different nations	*0.71*	0.23	0.07	-0.02
21. The capability of the different national teams	*0.71*	-0.07	0.47	-0.08
13. Previous championship records *	*0.69*	0.03	0.13	0.23
10. Own favourite national teams	0.49	0.22	0.08	0.07
17. The population of the different nations	-0.04	*0.74*	0.15	0.19
4. Number of soccer players in the different nations *	0.20	*0.73*	0.03	0.22
19. The geographic position of the nations	0.04	*0.71*	0.17	-0.16
3. The climate in the different nations	0.16	*0.70*	0.10	-0.17
12. Economic prosperity in the different nations	0.09	*0.68*	-0.02	0.16
18. The length of the soccer seasons in the nations	-0.04	*0.60*	0.57	0.11
6. Number of soccer clubs in the different nations *	0.27	*0.51*	0.14	0.36
11. The soccer interest in the different nations	0.49	0.49	-0.07	0.06
16. The playing system of the different national teams	0.16	0.21	*0.84*	0.00
1. Wins / Number of matches played in the preliminary competition 2000-2001 *	0.20	-0.05	*0.71*	0.22
20. The competence and know-how of the coaches in the national teams	0.38	0.09	*0.69*	-0.08
15. Earlier matches between the national teams	0.26	0.33	*0.56*	0.22
14. Odds *	-0.08	0.14	0.11	*0.80*
5. World rankings *	0.44	0.07	0.08	*0.61*

Note. Variable numbers refer to the order in the measurement. The variables with numbers in italics were included in the factors. The numbers in parenthesis refer to explained variance. The asterisk denotes cues that were available in the form with information.

The ratings of the 21 items of information give additional insights into the judgmental processes of the participants. To examine how these variables were related, an exploratory factor analysis (principal components) was performed and showed that the variables could be condensed into four factors accounting for 59% of the total variance. The varimax rotated factor pattern is exhibited in Table 4 and could be interpreted as follows.[5] The first factor covered general soccer information linked with the national teams, while the second factor predominantly involved facts about the countries. The third factor included specific details related to soccer like playing system, results in the preliminary competition, capability, and track records. The fourth factor was based on two related variables: odds and world ranking. In a sense, these variables represent a condensed form of information, because when computing odds and rankings in real life, consideration is taken to other items of information.

The mean ratings of the cues included in each informational factor were calculated. As shown in Table 5, the Swedish knowledgeable students as well as the experts reported to have relied on general and specific facts about soccer (i.e., factor 1 and 3) to a larger extent compared to the naïve participants. This tendency was found to be significant (MS = 33.99 and 17.62, $F(3, 250) = 6.56$ and 22.01, $p < 0.000$). Given the low ratings of the second factor, the participants appeared to have paid very little attention to country-specific details. Interestingly, the experts seemed to have put little weight to odds and rankings, whereas the naïve Swedish students with information were inclined to look at these variables (factor 4). Furthermore, some experts and knowledgeable Swedish students added cues they claimed to have considered when making their forecasts.

To investigate whether the naïve students differed from those with knowledge, four ANOVA-models were run. These models had two between-subjects variables (access to information and the three levels of knowledge) and the interaction of those variables. As they had not been exposed to any experimental manipulation, the experts were excluded from these analyses.[6] Significant between-subjects effect concerning knowledge was found with respect to general and specific soccer facts (MS = 88.90 and 36.25, $F(2, 197) = 59.61$ and 23. 45, $p < 0.000$) but not with respect to odds and rankings. In other words, those who knew more looked at more soccer-related information. As regards access to information, significant between-subjects effect was observed for general soccer facts and odds and rankings (MS =

[5] Many thanks to Dr. Alastair McClelland for suggesting interpretations of the results of the factors analysis.
[6] As earlier mentioned, the sample was randomly divided in order to manage the problem with unequal observations. The reported statistics are based on models including the whole sample.

6.72 and 32.58, F (1, 197) = 4.51 and 13.16, p < 0.04 and 0.000). Thus, the informed participants seemed to have relied on some items of the given information. The only significant interaction effect was linked with odds and rankings (MS = 12.28, F (2, 197) = 4.96, p < 0.008). It was, in particular, the naïve Swedish students with access to information, who considered odds and rankings.

Table 5. *Mean ratings of the four informational factors and average response time.*

	Factor 1: General soccer facts Mean (SD)	**Factor 2: Country specific facts Mean (SD)**	**Factor 3: Specific soccer facts Mean (SD)**	**Factor 4: Odds and ratings Mean (SD)**	**Response time in minutes Mean (SD)**
Experts	4.47 a (0.80)	0.99 (0.82)	2.42 a (1.38)	1.88 a (1.29)	16.30 a (11.41)
Knowledgeable Swedish students with information	4.02 b (1.20)	1.04 (1.24)	2.22 b (1.68)	2.53 b (1.32)	11.73 (3.84)
Knowledgeable Swedish students with no information	4.67 c (0.76)	1.17 (1.04)	2.58 c (1.32)	2.67 c (1.57)	12.81 (4.79)
Naïve Swedish students with information	2.81 ab (1.14) cd	0.71 (0.65)	1.12 ac (1.09)	3.01 a (1.73)	12.16 (3.63)
Naïve Swedish students with no information	3.63 ace (1.37)	1.04 (1.00)	1.14 ac (1.04)	1.96 (1.48)	10.30 a (3.36)
American students with information	1.67 abc (1.67) def	1.04 1.05)	1.48 (1.11)	2.73 (2.03)	13.14 (3.92)
American students with no information	1.42 abc (1.69) def	1.03 (1.29)	0.74 ab (1.19) c	0.94 bc (1.48)	11.83 (4.25)

Note. The factors were constructed in accordance to the result of the rotated factor pattern as exhibited in Table 4. The factors represent the average weight of different cues, which were rated on a six-point scale. The higher the value, the greater weight was taken to the individual factor.
The letters *a, b, c, d, e,* and *f* denote participant groups that differed significantly (p<0.05) from each other as shown by Tamhane's post-hoc tests. For instance, in the first column the letter *a* indicates that the experts differed significantly from the naïve Swedish students with and without information as well as American students with and without information. In the third column *a* indicates that there were significant differences between experts and the naïve Swedish students with information.

Response time gives further, but rough, insights into judgmental processes. In comparison to the others, the experts claimed to have taken more time answering the questionnaire (see Table 5). Response time was negatively correlated with number of correct forecasts (r = -0.17). The 33 participants who spent more than 18 minutes responding to the questionnaire made fewer correct forecasts (M = 8.85, SD = 1.40) than the 30 participants (M = 9.80, SD = 1.50) taking less than eight minutes (t (61) = 2.61, p < 0.01). Surprisingly, those two groups did not differ with respect to self-reported information use and confidence.

Finally, in a further attempt to evaluate the judgmental processes, participants with the poorest and the best predictions were compared. Analyses showed that there were no significant differences with regard to judgmental processes between those who had less than eight correct forecasts (n=15) and those who had more than eleven correct forecasts (n=11).

5. Conclusion and Discussion

In the introduction several questions were raised. Those questions are answered below.

How good are experts at forecasting soccer? Our study finds that the experts made poor forecasts of the first round of the World Cup in soccer. Specifically, the average expert managed to predict slightly better than chance. The evidence is consistent with Sherden (1998) who noted that experts in several different areas (e.g., weather, economic, and population) often fail to produce accurate prognoses. Consistent with earlier research (Ayton, 1992; Bolger & Wright, 1992; Allwood & Granhag, 1999), our study also finds that experts were very confident and overestimated the accuracy of their forecasts. This tendency of the soccer experts being overconfident is not surprising. As experts, they must mediate the image of being accurate regardless of their true performance. Few would otherwise pay any attention to their predictions of outcomes of sports events.

Are forecasts of experts better than those made by non-experts? Regardless of whether they had access to information or not, the participants with limited knowledge were able to produce slightly better predictions than the experts. Although the predictions were not significantly better, one may safely conclude that experts do not outperform non-experts with respect to forecasting soccer. This result may seem remarkable, but corresponds with earlier research (Camerer & Johnson, 1991).

How do experts and non-experts process information when forecasting soccer? Our data suggests that experts and knowledgeable participants put more weights to soccer-related facts than the naïve participants, who tended to rely on less information. The experts also reported that their judgments were based on extensive search and analysis of information. Regardless of whether information was present or absent, the naïve Swedish students claimed that they chose the teams they had hear of earlier with regard to soccer. In other words, they seemed to have employed the recognition heuristic. Unfortunately, no data were available on the judgmental processes of the American students who received no information. Because there was a positive relationship between their forecasts and their team knowledge, one might speculate that they used the recognition heuristic.[7] They might have based their forecasts on their knowledge of countries. For instance, they might have picked the countries they assumed were the biggest trade partners to the US. Irrespective of their judgmental processes and despite their limited knowledge, the American students performed, on average, surprisingly well.

As regards the other two types of fast and frugal heuristics (i.e., take-the-best and the minimalist heuristics), no support was found that they were explicitly in use. This finding relates to Newell & Shanks (2003) who reported that many subjects in their experiments rejected the "take-the-best" heuristic despite the presence of conditions promoting its use.

Caution must be taken to the present study's findings on information-processing strategies, as they are based on self-reported behavior. For instance, it could be the case that the experts stated to have considered more cues than they really did in order to comply with their reputation of having expertise. Although self-reported behavior might have made the measurements less reliable, it should be noted that they, nevertheless, pointed towards similar results.

Are there any information-processing strategies that successfully simplify forecasting soccer? We could not find that the participants explicitly employed a strategy that resulted in more accurate forecasts. However, a model based on the average team knowledge by the naïve Swedish students managed to predict 12 teams correctly. This model outperformed the experts and the non-experts; a finding consistent with the literature (Camerer & Johnson, 1991). A simulation model taking numerous variables into account predicted 11 winning teams correctly (O'Donoghue, 2002). Thus, we argue that a recognition-based model may be appropriate to employ when forecasting the outcome of soccer events.

In conclusion, the present paper has shown that soccer experts do not produce more accurate forecasts than do uninformed, naïve people and that a simple recognition-based heuristic appears to be highly adequate to use when forecasting the World Cup in soccer. It should be emphasized that the present paper dealt with the harmless area of predicting outcomes of sports events. To what extent these findings could be generalized to other areas of judgmental forecasting remains to be seen.

[7] In fact, such a relationship existed also for the other categories of participants. By definition, the experts and the knowledgeable students could not have employed the recognition heuristic: They simply knew too much about the teams. Recall that the recognition heuristic demands binary knowledge; either a team is recognized or it is not (Goldstein & Gigerenzer, 2002).

References

Allwood, C. M., & Granhag, P. A. (1999). Feelings of confidence and the realism of confidence in everyday life. In P. Juslin & H. Montgomery (Eds.), *Judgment and decision making: Neo-Brunswikian and process-tracing approaches*. Mahwah, New Jersey: Lawrence Erlbaum Associates, pp. 123-146.

Andersson, P. (In press). Does experience matter in lending? A process-tracing study on experienced loan officers' and novices' decision behavior. Forthcoming in *Journal of Economic Psychology.*

Aukutsionek, S. A., & Belianin, A. V. (2001). Quality of forecasts and business performance: A survey study of Russian managers. *Journal of Economic Psychology*, 22, 661-692.

Ayton, P. (1992). On the competence and incompetence of experts. In F. Bolger & G. Wright (Eds.), *Expertise and decision support*. New York: Plenum Press, pp. 77-105.

Ayton, P. (1998). How bad is human judgment? In G. Wright & P. Goodwin (Eds.), *Forecasting with judgment*. Chichester, UK: Wiley, pp. 237-267.

Ayton, P., & Önkal, D. (1997). *Forecasting Football Fixtures: Confidence and Judged Proportion Correct*. Unpublished manuscript, Department of Psychology, City University, London.

Barber, B. M., & Odean, T. (2002). Online investors: Do the slow die first? *The Review of Financial Studies, 15*, 455-487.

Bolger, F., & Wright, G. (1992). Reliability and validity in expert judgment. In F. Bolger & G. Wright (Eds.), *Expertise and decision support*. New York: Plenum Press, pp. 47-76.

Borges, B., Goldstein, D. G., Ortmann, A., & Gigerenzer, G. (1999). Can ignorance beat the stock market? In G. Gigerenzer, P. M. Todd, & the ABC group (Eds.), *Simple heuristics that make us smart*. New York: Oxford University Press, pp. 59-74.

Boulier, B. L., & Stekler, H. O. (1999). Are sports seedings good predictors?: an evaluation. *International Journal of Forecasting, 15*, 83-91.

Boulier, B. L., & Stekler, H. O. (In press). Predicting the outcomes of National Football League games. Forthcoming in *International Journal of Forecasting*.

Camerer, C. F., & Johnson, E. J. (1991). The process-performance paradox in expert judgment: How can experts know so much and predict so badly? In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*. New York: Cambridge Press, pp. 195-217.

Conroy, R., & Harris, R. (1987). Consensus forecasts of corporate earnings: Analysts' forecasts and time series methods. *Management Science, 33*, 725-738.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, & ABC Group (Eds.), *Simple heuristics that make us smart*. New York: Oxford University Press, pp. 97-118.

De Bondt, W. F. M. (1991). What do economists know about the stock market? *The Journal of Portfolio Management*, 85-91.

Forrest, D., & Simmons, R. (2000). Forecasting sport: the behaviour and performance of football tipsters. *International Journal of Forecasting, 16*, 317-331.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650-669.

Gigerenzer, G., Todd, P. M., & ABC Group. (1999). *Simple heuristics that make us smart*, New York: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109*, 75-90.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge Press.

Ladouceur, R., Giroux, I., & Jacques, C. (1998). Winning on the horses: How much strategy and knowledge are needed? *The Journal of Psychology, 132*, 133-142.

Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision-making work? In G. Gigerenzer, P. M. Todd, & ABC Group (Eds.), *Simple heuristics that make us smart*. New York: Oxford University Press, pp. 119-140.

Mills, T. C., & Pepper, G. T. (1999). Assessing the forecasts: an analysis of forecasting records of the Treasury, the London Business School and the National Institute. *International Journal of Forecasting, 15*, 247-257.

Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision-making. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 53-65.

O'Donoghue, P. (2002). *Simulator prediction of 2002 FIFA World Cup*. Unpublished manuscript, School of Applied Medical Sciences & Sports Studies, University of Ulster, Ulster, Northern Ireland.

Oskamp, S. (1982). Overconfidence in case-study judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press, pp. 287-293.

Russo, J. E., & Schoemaker, P. J. H. (1992). Managing overconfidence. *Sloan Management Review, Winter*, 7-17.

Sherden, W. A. (1998). *The fortune sellers: The big business of buying and selling predictions*, New York: Wiley.

Sundali, J. A., & Atkins, A. B. (1994). Expertise in investment analysis: Fact or fiction. Organizational Behavior and Human Decision Processes, 59, 223-241.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

Appendix A. *Descriptive statistics of the responses to the questions with regard to how well the participants knew the different teams.*

Country (World rankings)	**Soccer experts**		**Knowledgeable Swedish students**		**Naïve Swedish students**		**American students**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Group A								
France (1)	5.54	0.64	5.20	0.94	2.76	1.88	0.91	1.63
Senegal (44)	1.87	1.75	1.16	1.42	0.38	0.95	0.06	0.24
Uruguay (22)	2.58	1.55	1.84	1.53	0.50	0.88	0.45	1.09
Denmark (19)	4.83	1.00	3.88	1.25	1.77	1.64	0.39	1.06
Group B								
Spain (7)	4.85	1.21	4.43	1.33	1.99	1.74	0.64	1.43
Slovenia (28)	2.62	1.54	1.66	1.31	0.36	0.72	0.27	0.80
South-Africa (36)	2.48	1.46	1.36	1.20	0.53	0.84	0.42	0.87
Paraguay (15)	2.90	1.54	2.00	1.61	0.55	0.95	0.33	0.82
Group C								
Brazil (3)	4.90	1.03	4.76	1.13	2.73	1.89	0.91	1.74
China (52)	1.69	1.49	1.09	1.15	0.34	0.62	0.33	0.78
Turkey (25)	3.87	1.58	2.93	1.62	0.92	1.18	0.42	1.03
Costa Rica (27)	1.81	1.40	1.47	1.25	0.48	0.82	0.45	0.87
Group D								
Korea (41)	1.61	1.25	1.32	1.24	0.42	0.85	0.36	0.86
Poland (33)	3.00	1.39	1.99	1.44	0.42	0.76	0.52	0.94
USA (13)	2.60	1.36	2.31	1.22	0.99	1.28	1.30	1.85
Portugal (6)	4.42	1.16	4.07	1.38	1.32	1.54	0.70	1.40
Group E								
Germany (10)	4.96	1.04	4.34	1.16	2.20	1.72	0.64	1.25
Saudi-Arabia (35)	1.46	1.20	1.15	1.22	0.48	0.95	0.33	0.78
Ireland (18)	4.04	1.22	2.69	1.69	0.86	1.16	0.55	1.25
Cameroon (19)	3.27	1.61	2.70	1.59	1.33	1.48	0.30	0.64
Group F								
Argentina (2)	4.88	1.28	4.65	1.42	2.07	1.86	0.70	1.53
Nigeria (32)	3.90	1.45	3.43	1.49	1.27	1.53	0.33	0.89
England (12)	5.58	0.54	5.47	0.71	3.42	1.69	0.85	1.75
Sweden (16)	5.88	0.47	5.78	0.56	4.08	1.69	0.52	1.00
Group G								
Italy (4)	5.40	0.82	4.95	1.06	2.67	1.80	0.97	1.90
Ecuador (37)	1.85	1.42	1.24	1.31	0.47	0.94	0.30	0.68
Croatia (23)	3.37	1.61	2.50	1.53	0.55	0.97	0.36	0.86
Mexico (8)	2.62	1.62	2.26	1.27	0.75	1.18	0.76	1.41
Group H								
Japan (38)	2.40	1.49	2.19	1.32	0.67	0.85	0.61	1.34
Belgium (21)	3.58	1.41	2.49	1.46	0.80	1.17	0.30	0.81
Russia (24)	2.90	1.55	2.05	1.31	0.72	1.23	0.61	1.12
Tunisia (29)	1.71	1.54	1.00	1.03	0.26	0.66	0.24	0.61

Note. Before making the forecasts and the experimental manipulation, the participants responded to the questions concerning how well they knew the teams qualified to the World Cup in soccer. Consequently, there was no need to differentiate the participant groups with respect to whether they had obtained information or not. The countries are reported in their respective playing groups. Knowledge of each team was rated on a seven-point verbally anchored scale ranging from the “not at all” (0) to “very well” (6). World rankings are reported as they were before the start of the World Cup. World rankings correlated strongly with team knowledge ($-0.72 < r_s < -0.81$), implying that the participants had better (lesser) knowledge about the teams that had top (low) rankings.

Appendix B. *Percentage of participants forecasting the advancement of each team.*

Country (World rankings)	Soccer experts	Knowledgeable Swedish students		Naïve Swedish students		American students	
		With information	With no information	With information	With no information	With information	With no information
Group A							
France (1)	100	100	100	97	100	100	94
Senegal (44) *	13	3	7	11	9	7	6
Uruguay (22)	19	30	45	50	30	47	50
Denmark (19) *	67	67	48	42	61	47	50
Group B							
Spain (7) *	100	100	100	95	91	100	100
Slovenia (28)	27	20	14	13	15	7	17
South-Africa (36)	2	20	11	18	43	13	39
Paraguay (15) *	71	60	75	74	52	80	44
Group C							
Brazil (3) *	100	100	100	100	100	100	94
China (52)	17	13	7	13	15	13	39
Turkey (25) *	77	83	86	82	74	67	33
Costa Rica (27)	6	3	7	5	11	20	33
Group D							
Korea (41) *	35	7	18	21	31	13	44
Poland (33)	60	53	48	29	19	53	11
USA (13) *	10	40	34	55	72	60	83
Portugal (6)	96	100	100	95	78	73	61
Group E							
Germany (10) *	94	97	100	100	96	93	94
Saudi-Arabia (35)	2	3	7	8	6	7	11
Ireland (18) *	29	53	23	21	31	47	78
Cameroon (19)	75	47	70	71	67	53	17
Group F							
Argentina (2)	98	93	95	89	76	100	83
Nigeria (32)	17	3	25	5	13	0	17
England (12) *	58	63	50	66	81	93	67
Sweden (16) *	27	40	30	39	30	13	33
Group G							
Italy (4) *	100	100	100	100	100	93	67
Ecuador (37)	4	7	9	3	20	0	44
Croatia (23)	71	60	68	50	39	40	39
Mexico (8) *	25	33	23	47	41	60	50
Group H							
Japan (38) *	63	47	70	37	54	27	67
Belgium (21) *	52	63	50	68	65	80	50
Russia (24)	85	73	64	87	65	93	67
Tunisia (29)	0	17	16	8	17	0	17

Note. Asterisk denotes actual advancement to the second round. Because the participants selected two teams in each playing group, the percentage numbers with regard to the playing groups add up to 200 %. World rankings are reported as they were before the start of the World Cup in soccer.