Judgment and Decision Making, Vol. 6, No. 5, July 2011, pp. 381–391

Effects of ignorance and information on judgments and decisions

Peter Ayton* Dilek Önkal† Lisa McReynolds‡

Abstract

We compared Turkish and English students' soccer forecasting for English soccer matches. Although the Turkish students knew very little about English soccer, they selected teams on the basis of familiarity with the team (or its identified city); their prediction success was surprisingly similar to knowledgeable English students—consistent with Goldstein and Gigerenzer's (1999; 2002) characterization of the recognition heuristic. The Turkish students made forecasts for some of the matches with additional information—the half-time scores. In this and a further study, where British students predicting matches for foreign teams could choose whether or not to use half-time information, we found that predictions that could be made by recognition alone were influenced by the half-time information. We consider the implications of these findings in the context of Goldstein and Gigerenzer's (2002, p. 82) suggestion that "... no other information can reverse the choice determined by recognition" and a recent more qualified statement (Gigerenzer & Goldstein, 2011) indicating that two processes, recognition and evaluation guide the adaptive selection of the recognition heuristic.

Keywords: recognition heuristic, judgmental forecasting, inference, heuristics.

1 Introduction

How does knowledge affect judgment and decision-making? While much judgment research has considered the role of expertise, Goldstein and Gigerenzer (2002) have made the startling claim that, under certain circumstances, ignorance can benefit judges attempting to make inferences from their knowledge. For example, Goldstein and Gigerenzer (2002) asked both American and German students which is the bigger city: San Antonio or San Diego? Sixty-two percent of the Americans correctly named San Diego—but 100% of the German students were correct. All of the German students had heard of San Diego but only about half had heard of San Antonio. Half of the German students would therefore be able to apply a recognition heuristic—if you recognise one and not the other, pick the city you recognise. As people usually hear about the bigger cities of foreign countries before the smaller ones—i.e., recognition correlates with the criterion being predicted—this cue will have some validity. Because the American students had heard of both cities they couldn't utilise this cue and had to rely on other, apparently less valid, cues. In terms of accuracy, it seems that, when it comes to using knowledge to make inferences, less can sometimes mean more. Specifically, when the validity of the recognition cue exceeds the validity of the knowledge that can be applied when both items are recognised, a "less is more" effect is predicted.

To corroborate this "less is more" effect, Goldstein and Gigerenzer (2002) administered two quizzes to 52 University of Chicago students—one about the biggest 22 German cities and one about the biggest 22 American cities. Each quiz was comprised of 100 randomly selected pairs of the biggest 22 cities for the country. They found that the American students were slightly more accurate about German cities than American cities. Furthermore, Goldstein and Gigerenzer (2002) cited Hoffrage (1995; see also Hoffrage, 2011) who apparently found that German students were slightly more accurate (albeit non-significantly) at making decisions about American cities than about German cities. Despite a lifetime of experience and learning about the cities in one's native country it seems that one can be more accurate when making decisions about the cities in a country that is considerably less familiar. Since these initial experiments, further research (Beaman, Smith, Frosch, & McCloy, 2010; Hoffrage, 2011; Katsikopoulus, 2010; Smithson, 2010) has extended our understanding of the conditions when a less-is-more effect may occur.

In another experiment to confirm that the recognition heuristic predicted inferences Goldstein and Gigerenzer (2002) found that subjects adhered to the predictions of the recognition heuristic on 90% of the occasions that it was applicable and concluded that the recognition heuristic captured the vast majority of inferences. The more

We are grateful to Ben Newell, Ulrich Hoffrage, Julian Marewski, Dan Goldstein and Rüdiger Pohl for their helpful comments on an earlier draft of this manuscript.

*Department of Psychology, City University London EC1V 0HB, United Kingdom E-mail: P.Ayton@city.ac.uk.

†Bilkent University

‡City University London

general theoretical point that Goldstein and Gigerenzer (2002) drew from these demonstrations is that a very simple one-reason heuristic using very limited knowledge can make surprisingly accurate decisions. Even ignorance can sometimes be helpful because a simple mental heuristic like the recognition heuristic operates so as to take advantage of the structure of information to make good inferences. In particular, the recognition heuristic is successful when ignorance, specifically a lack of recognition, is systematically rather than randomly distributed, that is, when it is strongly correlated with the criterion.

Of particular interest to us here is Goldstein and Gigerenzer's (2002) suggestion that the recognition heuristic uses a noncompensatory rule. Even when other information about a recognized alternative is available, it never overrides the weight placed on simple recognition: "If one object is recognized and the other is not, then the inference is determined; no other information about the recognized object is searched for and, therefore, no other information can reverse the choice determined by recognition" (Goldstein & Gigerenzer, 2002, p. 82).[1]

Consistent with their suggestion that the recognition heuristic utilises a non-compensatory decision rule, Goldstein and Gigerenzer (2002) reported evidence that use of the recognition cue was unaffected by the presence of other conflicting information. We give a more detailed account of their study in the general discussion but, in short, they taught their respondents that cities with soccer teams were bigger than those without in 78% of possible city pairings. Yet, when confronted with a choice between a city they recognised and knew did not have a soccer team and one they did not recognise but that might have a soccer team, respondents chose the recognised city in 92% of the pairs. Thus, additional diagnostic information about soccer teams and city sizes had no measurable impact on decisions where the recognition heuristic was applicable.

A number of subsequently published studies clearly show that recognition, at least in a general sense, plays a role in prediction (e.g. Andersson, Edman & Ekman, 2005; Reimer & Katsikopoulos, 2004; Serwe & Frings, 2006) including for soccer matches (Pachur & Biele, 2007). These studies did not test the non-compensatory use of recognition[2]—critical for distinguishing the recognition heuristic from similar but distinct proposals such as the availability heuristic (Tversky & Kahneman, 1973), which attribute judgment to the relative ease of retrieval, or the fluency heuristic (Schooler & Hertwig, 2005), which assumes the most fluently processed item has the highest criterion value.

Nevertheless other studies have challenged the non-compensatory use of recognition. For example Newell and Shanks (2004) found in two cue-learning experiments that people did not place any special status on recognition. Most of their subjects learned to use recognition-based information when it was a good predictor and to essentially ignore it when it was a poor predictor. They found little evidence that recognition is treated any differently from other cues in the environment.

Note that, although these results show that laboratory-induced recognition may be overruled by "given" information (supplied by the experimenter), they do not show that natural (pre-experimental) recognition can be overruled by information drawn from memory. In their recent review of research into the recognition heuristic Gigerenzer and Goldstein (2011) discuss what they identify as "Misunderstanding #3" (pp. 107–108) and emphasise that studies focusing on inferences from given information or experimentally induced recognition are beyond the domain of the recognition heuristic.[3]

Newell and Shanks's (2004) study and others that have tested the non-compensatory use of recognition (Oppenheimer, 2003; Newell & Fernandez, 2006; Bröder & Eichler, 2006; Pohl, 2006; Richter & Späth, 2006) have been criticised by Pachur, Bröder & Marewski (2008) who argued that most of the studies testing the non-compensatory use of recognition used experimental situations that differ in potentially critical ways from the situations for which Goldstein and Gigerenzer (2002) formulated the heuristic. According to Pachur et al. (2008, and also Pachur & Hertwig, 2006) an ideal test of the recognition heuristic uses natural cue knowledge, rather than teaching it in the same laboratory setting in which choices are elicited—a failing of even Goldstein and Gigerenzer's test of non-compensatory cue use—and natural (pre-experimental) recognition—a failing of Newell and Shanks's two cue-learning experiments. Moreover, according to Pachur et al., research should test cases where recognition is a good predictor of the criterion and also preclude the possibility that knowledge other than cue knowledge could inform decisions—conclusive criterion knowledge (e.g., knowing that a recognized city is very small and so guessing that another unrecognized city is bigger) could allow decision makers to make judgments without engaging in the inductive inference that Goldstein and were theorizing about.

[1]In Goldstein and Gigerenzer's (2011) discussion of "Misunderstanding #2" (p. 107) they clarify their position that the Recognition heuristic ignores cues but not information in general, such as information concerning criterion values or the recognition validity (but see Newell's [2011] doubts about this distinction).

[2]Pachur and Biele's study included evaluation of a compensatory model (TALLY) but did not test if additional information compensated recognition.

[3]Goldstein and Gigerenzer's (2002) paper reported experimental tests of the recognition heuristic that used laboratory-induced recognition and information from "givens" (with no caveats signalling any theoretical improprieties in so-doing) which may have contributed to Misunderstanding #3.

The experiments presented here investigate how people use recognition and are exempt from the issues raised by Pachur et al.—although, like Goldstein and Gigerenzer (2002), we do test the impact of information from "givens", namely presented information rather than information drawn from memory. While this issue was not referred to by Goldstein and Gigerenzer (2002), it was mentioned in Gigerenzer and Goldstein (1996), which first introduced the recognition heuristic. It is further emphasized by Gigerenzer and Goldstein (2011), who identify the presumption that the recognition heuristic applies to inferences from "givens" as a misconception and describe tests of the recognition heuristic using inferences from given information as going "beyond the domain of the recognition heuristic". Although their own test of the impact of conflicting information tested subjects' use of given information and an experimenter-supplied cue validity (Goldstein & Gigerenzer, 2002) it did not provide what Gigerenzer and Goldstein (2011) describe as cue information about unrecognized objects—which may be why Goldstein and Gigerenzer (2002) made no reference to the notion that their own test was beyond the domain of the recognition heuristic. Indeed they concluded: "This result supports the hypothesis that the recognition heuristic was applied in a noncompensatory way" (p. 83).

In our first study we examined the judgmental predictions of two different groups of subjects who, like the American and German students studied by Goldstein and Gigerenzer, would be expected to have different knowledge of a domain. This study—conducted in 1995, long before we read about the recognition heuristic—was cited by Goldstein and Gigerenzer (2002) as, rather fortuitously we must admit, it illustrated the surprising benefit of ignorance for judgment. We studied Turkish and English students' predictions of English soccer matches. We expected the Turkish subjects to have very little knowledge of the domain while the British subjects would have a good deal of knowledge. Although predicting soccer matches is not quite the same as making inferences about city sizes, we were able to test the accuracy of subjects' predictive inferences by comparing the forecasts with the outcomes of the games. Given the Turkish subjects' very limited knowledge of English soccer, we envisaged that relative familiarity with the names of the English cities and towns that often make up the names of soccer clubs might very often be their only basis for making forecasts. So, although we did not have the recognition heuristic in mind when we designed this study, we measured subjects' familiarity with city names—including those for which they had no familiarity at all—in order to examine the relationship between this variable and their forecasts. We were also able to study the effect of providing additional diagnostic information on the Turkish subjects' judgments. After they made their initial forecasts we provided them with the half-time scores of the matches to investigate whether their forecasts with additional information reflected any integration of the additional information in their judgment. We note here that, as we discuss in more detail below, an objection to this study is that it goes beyond the domain of the recognition heuristic because, in providing the half-time scores, we provide what Gigerenzer & Goldstein (2011) describe as cue information about unrecognized objects (Gigerenzer & Goldstein, 2011)—which Goldstein and Gigerenzer (2002) did not do.

2 Experiment 1

We tested the abilities of subjects to make forecasts about the 32 matches that comprise the third round of the English Football Association trophy (F.A. cup).

2.1 Method

All subjects were presented with a list of the 32 soccer matches drawn for the 1994 English F.A. cup 3rd round. The F.A. cup is the major knockout tournament for English soccer clubs. Many hundreds of soccer teams enter this tournament and compete through preliminary rounds to reach the final stages of the competition; the draw for the 3rd round is the first occasion when the country's top professional teams enter the contest. As teams are randomly paired, famous clubs can be drawn against less well-known lower clubs and occasionally even very obscure amateur teams who have managed to reach the latter stages.

Subjects were instructed to select the team that they thought would win each match and then rate their confidence in their forecast on a 50–100% scale. After making judgments for all the matches they were asked to state how many of their 32 forecasts they expected to be correct. Before making their forecasts, the Turkish subjects completed a questionnaire where for each of the mentioned teams they rated their degree of familiarity with the name of the team on a 7-point scale (where 1=no familiarity at all, 7= very high familiarity). Given their very limited knowledge of English soccer we hypothesised that relative familiarity with the names of the English towns might very often be their only basis for making forecasts.

As Turkish subjects made their forecasts after most of the games had actually taken place, we were able to present these subjects with the half-time scores of most of the matches to serve as an additional cue.[4] Accord-

[4] We assumed (safely as it turned out) that the Turkish subjects would not know the results of these matches and so could not use knowledge of results to make their predictions. The English fans were tested before the matches were played and we did not examine their reaction to

Table 1: Forecasting Performance (Experiment 1). Numbers are group medians of subjects' mean scores.

Forecasts	Proportion correct	Judged proportion correct	Mean confidence	Overconfidence
Turkish	62.5%	62.5%	67.2%	4.5%
Turkish (with half-time scores)	60.0%	68.8%	69.5%	9.0%
British	65.6%	56.3%	70.1%	5.2%

ingly the Turkish subjects then repeated their forecasts for the matches with the additional half-time score cue and again rated their confidence in their forecast on a 50–100% scale.

54 British students and 50 Turkish students made forecasts for all 32 English F.A. cup 3rd round soccer matches. The British subjects were recruited contingent on a stated interest in soccer and successful completion of a short quiz that probed knowledge of the outcome of the previous season—respondents not interested in soccer or not successfully completing the quiz were not recruited. The Turkish subjects were recruited without these requirements. Following their predictions for all the matches the Turkish subjects gave predictions for 19 of the matches with the half-time results. We removed three Turkish respondents who often rated their forecast confidence less than .50. Other data were missing due to some subjects omitting some familiarity ratings and/or predictions.

2.2 Results

In developing the case for the recognition heuristic Goldstein and Gigerenzer (1999, 2002) discussed our experimental data—previously only available in an unpublished manuscript (Ayton & Önkal, 1997). As they reported, we found that, despite very limited knowledge, the Turkish students had a surprising ability to forecast the English soccer matches. Table 1 shows that the Turkish students correctly forecast 62.5% of the matches; although inferior to the 65.6% success rate of the British students (Mann-Whitney U=1053; p=.052), the difference is small and not (quite) statistically significant.

The familiarity ratings confirm our assumption that the Turkish subjects were largely ignorant about the English soccer teams. In 82% of the 1036 total cases (subjects x matches), Turkish subjects had no familiarity at all with either just one (n=662) or both (n=184) of the UK teams (or the city that often makes up part of English soccer team names)—many also spontaneously protested their ignorance during the testing.

additional information.

To test for the application of the recognition heuristic, Goldstein and Gigerenzer (2002) examined subjects' decisions where one item was recognised and the other item was not recognised at all. Accordingly, we similarly studied the association between rated familiarity and forecasts. Table 2 presents the data from our experiment—collated so that we can consider the impact of rated familiarity on forecasts for the 662 cases where subjects forecast a match involving one team (or the city name that often makes up part of English soccer team names—e.g. "Manchester United") that they indicated they were totally unfamiliar with and one team where they indicated at least some familiarity.

Table 2 shows that, when forecasting a match involving one team that was not at all familiar and another team that was rated as familiar to some degree, the Turkish students selected the familiar team on a very high proportion of occasions. As Goldstein and Gigerenzer (2002) reported, our data are consistent with the use of the recognition cue—among the pairs of soccer teams in which subjects rated one team as completely unfamiliar and the other as familiar to some degree, they chose the familiar team in 627 (95%) of 662 cases. Plainly, familiarity is an influential cue as well as an effective cue.

2.2.1 Impact of half-time scores

The performance of the Turkish respondents did not improve with the additional half-time score information (Table 1); although slightly lower, their proportion of correct forecasts with half-time information (60.0%) was not significantly different from their proportion correct before seeing the half-time scores (62.5%) (Wilcoxon signed ranks Z = −.81; p>.10).

Although there was no significant change in forecast accuracy, it is nonetheless possible that the half-time scores had some impact on the Turkish subjects' selections. Accordingly we collated the data so as to directly investigate whether respondents altered their forecasts when the half-time cue conflicted with the recognition cue. Table 3 shows the effect of the half-time information on subjects' forecasts for those (388) cases where both (a) the recognition heuristic was applicable and (b) half-time

Table 2: Subject mean forecasts for cases where recognition cue was applicable (Experiment 1)]

Familiarity rating of recognised team	Mean number of cases (Total cases)	Mean number of cases where recognised team is predicted to win (total cases)	Proportion of cases where recognised team is predicted to win	Mean confidence (when choosing recognised team)	Mean confidence (when choosing unrecognised team
2	4.94 (168)	4.56 (155)	88%	63%	62%
3	3.97 (155)	3.59 (140)	88%	67%	59%
4	3.90 (156)	3.83 (153)	99%	74%	65%
5	3.08 (80)	2.96 (77)	91%	74%	63%
6	2.25 (45)	2.25 (45)	100%	79%	--
7	3.93 (58)	3.93 (57)	100%	79%	--
2–7	14.39 (662)	13.65 (627)	94%	71%	62%

Table 3: Subjects' forecast winners of matches with half-time information (cases where recognition cue was used for initial forecast in Experiment 1). Numbers are frequencies of occurrence.

With ½ time scores forecast winner	Half-time score conflicts with recognition	Half-time score concurs with recognition	Half-time score Level
Recognised	25	117	222
Not recognised	17	0	7
Total	42	117	229

information was presented. There is a clear difference in the use of recognition when the half-time cue concurs or conflicts with recognition. When the half-time score favoured the familiar team, subjects *always* (117/117) selected the familiar team. However, when the half-time score favoured the unfamiliar team, respondents selected the familiar team on only 25/42 (59.5%) of occasions—a significantly lower proportion than 117/117 ($z=-7.28$, $p<.001$). When the half-time score was equivocal subjects selected the more familiar team on 222 of 229 occasions (97%). Combining the (229) games with equivocal half-time scores with those (117) where the half-time score was consistent with recognition creates a set of (346) cases when the half-time cue did not contradict recognition: subjects decisions were consistent with the recognition cue on 339/346 (98%) of occasions—significantly higher than the 25/42 (59.5%) of occasions when the two cues conflicted ($z=-9.77$, $p<.001$). Given the limited number of observations where the two cues were in conflict one may wonder how many of the respondents responded inconsistently with the recognition cue. Over a third (seventeen) of the 47 subjects made at least one decision inconsistent with recognition: on the 42 occasions when the cues conflicted 13 subjects made the 17 decisions inconsistent with recognition; when half-time scores were equivocal 7 subjects (including three of the 13 already noted) made the 7 (of 229) decisions inconsistent with recognition; when the half-time score conflicted with recognition. Although there is evidence for individual differences in people's use of recognition (Gigerenzer & Goldstein, 2011), our limited number of responses make it difficult to infer very much about this beyond the conclusion that choices inconsistent with recognition are not attributable to a maverick few.

Although there are relatively few observations where the recognition cue was antagonistic to the recognition cue, half-time information clearly had an influence on decisions that respondents could have adjudicated solely by reference to the recognition cue.

2.2.2 Confidence judgments

Table 1 shows that respondents were somewhat overconfident; while the judged proportion correct was not inappropriately high for the English or Turkish respondents without half-time information, mean confidence was higher than proportion correct, resulting in overconfidence. Table 2 shows that teams with higher rated familiarity are forecast with greater confidence. Across subjects, for those cases where respondents indicated they

were totally unfamiliar with one of the teams, mean confidence when choosing a recognised team was significantly higher than when choosing an unfamiliar team ($t(37)=4.26$; $p<.001$). Table 2 shows that, for those cases where respondents indicated they were totally unfamiliar with one of the teams, confidence increased as the familiarity of the recognised team increased; the mean of the individual subject correlations between confidence and familiarity for the other team is 0.381 ($p<.001$). This finding is consistent with research by Koehler (1996), which established that a simple model that associates a single strength value with each team can accurately account for the probability judgments of the outcomes of basketball matches made by basketball fans.

We also studied the effects of the half-time information on the confidence judgments. As shown in Table 2, mean confidence increased significantly ($Z=-4.49$, $p<.001$) with half-time information and, as performance was somewhat lower, overconfidence was substantially and significantly ($Z=-4.32$, $p<.001$) higher. We also conducted paired difference t-tests on those cases when the recognized team was predicted to win both before and after the half-time scores were given. Comparing the probabilities given for the same cases before and after the half-time scores were provided across all subjects we found that, when half-time scores concurred with recognition, forecast confidence increased significantly ($t(116)=7.27$, $p<.001$); but when the half-time scores conflicted with recognition, forecast confidence decreased significantly ($t(24)=4.54$, $p<.001$). Thus, not only are subjects' forecasts of which team will win affected by more than just recognition, but the confidence with which the forecasts are made also appears to be influenced in a compensatory fashion by other information.

Given Marewski, Gaissmaier, Schooler, Goldstein, and Gigerenzer's (2010) recent challenge to researchers to specify a model that out-predicts the recognition heuristic, and given that, despite the evidence for compensatory effects, the recognition heuristic predicted better than several alternative compensatory models, it would be interesting to conduct similar model comparisons. As these authors acknowledge it is possible that, although a person always chooses recognized alternatives over unrecognized ones, she still integrates knowledge into her decisions. For instance, rather than affecting decisions, such an integration of knowledge could result in reduced confidence.

2.3 Discussion

Our results show that decisions based on recognition of one team and ignorance of the other can be surprisingly effective; use of the recognition cue by the Turkish respondents was not only very frequently utilised—it also enabled them to achieve forecasting success comparable to the more knowledgeable English respondents. These findings exemplify the claim of proponents of the "fast-and-frugal" framework that so-called "one reason decision making" can be *ecologically rational*—i.e., simple heuristics can exploit structures of information in the environment (Gigerenzer, Todd, & the ABC Research Group, 1999). The forecasting success of the Turkish forecasters with very little knowledge is not so readily explicable by other approaches to decision making (see Yates, McDaniel & Brown's 1991 study showing experts making inferior predictions of stock prices to novices).

However, in contrast to Goldstein and Gigerenzer (2002), we did not find support for the notion that recognition-based decision-making is non-compensatory for given information. Indeed, we found that people do not rely exclusively on recognition when additional information is available. When presented with half-time scores, respondents used this information and, when it conflicted with recognition, significantly reduced their use of the recognition cue. This use of information not only affected the simple binary choices of respondents regarding which team would win but also the confidence with which they made their choices.

Several years ago, a reviewer of an earlier draft of this paper argued that this experiment did not properly test for recognition heuristic usage because it tests a case where, by definition, the recognition heuristic cannot be applied. According to this reviewer's argument, when the half-time scores are presented, there was information about both the recognised and unrecognised teams and therefore Goldstein and Gigerenzer's (2002) definition of "unrecognised" ("no knowledge of an object because one has never heard, smelled, touched, tasted, or seen it before" p. 77) is not met. (Gigerenzer & Goldstein, 2011, made the same point). Although it might be countered that recognition of the names would precede any evaluation of the half-time information, we decided to conduct a further experiment where recognition and the opportunity for consulting additional information were clearly segregated such that respondents could make a decision on the basis of recognition without any other knowledge being present, or, alternatively, choose to consult additional information. Eliciting only one forecast also reduces the possibility of recognition acquired in the experimental context affecting the results.

Our next experiment tests whether respondents search for additional information when one object is recognised and the other is not.

3 Experiment 2

This experiment utilised a single group of respondents who, like Goldstein and Gigerenzer's (2002) American students judging German cities, would have limited experience of the items they were asked to judge. We again asked respondents to rate the familiarity of a series of soccer teams and then forecast the winner of a series of knockout cup matches. In this experiment respondents had the option of seeking additional information (half-time scores) before making their predictions but were not able to inspect any additional information unless they explicitly sought it.

3.1 Method

3.1.1 Design materials and subjects

A single group of 17 subjects (16 females and 1 male) was recruited to take part in this experiment; all were undergraduate students at the City University in London. The soccer matches selected consisted of 44 knock-out cup matches from several different (non-U.K.) countries. We deliberately selected these so as to include soccer teams that would be unfamiliar to our respondents. The experiment used a computer program to present a series of soccer matches that we asked respondents to predict the winner of. All matches appeared individually in a random order generated differently for each respondent.

3.1.2 Procedure

The experiment took place in two phases—both presented via computer-controlled screens. In the first phase all subjects first rated the 88 team/city names for recognition. All the teams in this recognition test were presented singly in random order and respondents were instructed to rate the items from 1 (no familiarity with the team or the city or town that often makes up part of soccer team names name) to 7 (very familiar).

Following the rating of recognition the 44 matches with the two team names side by side were presented in random order. For each match subjects were instructed to mouse-click on the team they thought had won the match. All subjects made their forecasts after the games had actually taken place (they did not know the results) with the option of inspecting the half-time score. Displayed between the two team names was a graphic box titled "half-time score". If subjects chose to predict the match outcome without additional information then they could ignore the half-time score button, but, if they wished to use the half-time score to make their predictions, they could click on the box to reveal it.

After clicking on a team subjects then used a 50%-100% confidence scale to rate their confidence that their prediction was correct. The task was self paced.

3.2 Results and discussion

The familiarity ratings confirmed that our respondents had very limited knowledge of the events being forecast; 44.8% (670 of the 1496 [88 x 17]) of teams were rated as completely unfamiliar. Familiar (rating >1) and completely *un*familiar (rating =1) teams were paired together such that the recognition heuristic was applicable for 38.2% (286) of the 748 (44 matches x 17 respondents) forecast matches. Nonetheless, before forecasting the match winner, respondents actively sought additional information for a clear majority of these events; for those (286) cases where recognition could determine the issue respondents sought half-time information in 68.5% (196) of them. Given that there was no penalty associated with use of the half-time scores one might wonder why respondents did not consult this information more often. However, although there was no explicit cost for seeking the additional information there is plainly some cost in terms of time and effort to process this information, which may explain why respondents did not always seek it. With half-time scores respondents decided consistently with the recognition cue on 61.2% of occasions (120 times); the remaining 38.8% (76 times) they decided inconsistently with recognition. Where respondents did not seek half-time information (90 occasions) respondents used recognition on a similar proportion (61%, 55 times) of occasions.

Prediction accuracy was well above chance—73.8% of all (748) match predictions were correct. For those 286 cases where the recognition heuristic was applicable (one team was familiar another was totally unfamiliar) accuracy was 72.7%; for the 175 cases where the recognition heuristic was applicable and predictions were in line with it accuracy was 91.4%. Accuracy was below chance—43.2%—for those 111 cases where the heuristic was applicable but predictions went against recognition. For the (196) predictions that could be resolved by recognition but were also informed by half-time score information accuracy was 77.6%. A paired samples t-test comparing the performance of subjects on items where recognition was applicable when they inspected half-time scores against performance when they didn't inspect half-time scores showed a significant improvement in performance across individuals when they used half-time information ($t(16)=2.90$, $p=0.01$).

As with Experiment 1, we examined how half-time information affected subject forecasts—on those occasions when respondents chose to inspect it. Table 4 shows that when the half-time score concurred with the

Table 4: Subjects' forecast winners of matches with half-time information (cases where recognition cue was applicable and additional information was sought in Experiment 2). Numbers are frequencies of occurrence.

With ½ time scores forecast winner	Half-time score conflicts with recognition	Half-time score concurs with recognition	Half-time score Level
Recognised	2	38	80
Not Recognised	32	18	26
Total	34	56	106

recognition heuristic—i.e., favoured the same team—respondents used the recognition cue on most—38 of the 56—occasions; however, when the half-time score conflicted with the recognition heuristic—i.e., the two cues pointed in different directions—respondents used the recognition cue on just 2 of the 34 occasions. These two proportions 38/56 and 2/34 are significantly different from each other ($z=-8.34$, $p<.001$). Again we examined how many of the subjects eschewed recognition when the half-time score was in conflict with it: 15—nearly all—of our 17 subjects did this at least once, one subject never experienced an occasion when the two cues were in conflict and just one subject always utilised recognition in spite of (three) opportunities to use the half-time cue. Again it is clear that, even where recognition could determine the issue, half-time information influenced respondents' forecasting decisions. However, unlike Experiment 1, in this experiment the additional half-time information was available at respondents' discretion. Accordingly we conclude that this experiment demonstrates that, when one team was recognized and the other was not, then the inference was not always determined by recognition; on the contrary, additional information about the teams was sought on a majority of occasions and indeed, where this conflicted with recognition, it often reversed the choice that could have been determined by recognition. Notwithstanding Gigerenzer and Goldstein's (2011) recent qualifications, one could argue that this finding is inconsistent with Goldstein and Gigerenzer's (2002) original characterisation of the recognition heuristic and their finding that recognition is noncompensatory for given information—although one difference between our studies is that Goldstein and Gigerenzer (2002) did not provide what Gigerenzer and Goldstein (2011) call cue information about unrecognized objects and we did.

4 General discussion

Goldstein and Gigerenzer's (2002) test of noncompensatory inferences found that people do not compensate recognition with other information. Why do we find a compensating influence of conflicting information while Goldstein and Gigerenzer did not? One possibility is that Goldstein and Gigerenzer's test for compensatory effects was weaker than it appears. Below we argue that their mode of presenting the conflicting information may have undermined its impact and that the diagnosticity of the conflicting information was very low.

Goldstein and Gigerenzer's (2002) evidence came from an experiment using the German city sizes task conducted on American student subjects. They were informed that, after a training phase, they would be tested on pairs of cities drawn from the 30 largest in Germany. They instructed their subjects (as a recited fact) that, among the 30 biggest German cities, the 9 cities with soccer teams in the top league (the German "Bundesliga") are larger than the 21 without in 78% of all possible pairs. They then told subjects specific information about 8 "well known" cities drawn from the set of 30. Although the subjects believed that that they were sampling 8 cities at random, all subjects were given the same names of 4 "well known" cities with a soccer team and 4 "well-known" cities without a soccer team.[5]

After this training phase Goldstein and Gigerenzer (2002) presented their subjects with pairs of cities and asked them to select the bigger of each pair. The critical test focused on pairs where subjects recognised one city but knew it did not have a soccer team but did not recognize the other city (which may or may not have had a soccer team). The logic of this contrast is that, as the recognised city does not have a soccer team but the unrecognised one may, a subject who did not place any value at all on recognition should select the city which might have a soccer team over the one which they knew did not. Note that the role of recognition was never mentioned to the subjects and that all instruction concerned soccer teams.

In Goldstein and Gigerenzer's (2002) study the overwhelming majority of such decisions were consistent with the use of recognition. Twelve of 21 subjects made predictions entirely in accord with recognition and most other subjects deviated on only a few occasions. Overall 92% of the decisions made when conflicting cues were given were consistent with recognition.

[5]These were cities that they later indicated that they had heard of prior to the test.

While the test of the use of recognition when faced with conflicting information appears to impressively affirm the dominant power of recognition there are some grounds to doubt it. Firstly, there is a kind of conflict within the information presented during the training phase. While subjects were explicitly instructed that cities with soccer teams are bigger than those without on 78% of occasions, their experience with what they believed was a random sample of eight of the 30 biggest cities might well have created a rather different impression. They had studied eight cities during the training session and knew four that did not have soccer teams and four that did. These eight cities were selected by Goldstein and Gigerenzer because they (correctly) anticipated they would be recognised by American students—they were "well known" cities. Imagine then, in spite of being instructed that cities with soccer teams are bigger than those without on 78% of occasions, the effect of discovering that, among a putatively random selection of eight cities, all of which they happened to recognise and so would presumably assume are larger cities, half didn't have a soccer team. If having a soccer team was strongly diagnostic of city size one might expect a greater preponderance of cities one recognised to have soccer teams. In any event, whatever the respondents' expectations, the experience of discovering that half of the sample of eight well known—and therefore presumably large—cities did not have soccer teams might well have reduced the influence of the soccer team statistic below the level the explicit instructions were designed to invoke. Alternatively, given assurances that the soccer team information was correct and a $15 prize for correctness, respondents might have lowered their estimate of recognition validity, rather than the soccer team cue[6]—though which, if either, effect occurred remains unclear.

The Goldstein and Gigerenzer (2002) study on cue conflict thereby conflates two bits of information with two types of information. Subjects were *instructed* with a summary statistic about the validity of the soccer team as a cue to city size—but then *experienced* information that could undermine any sense they have of this.

A number of authors claim that there is a differential effect of experienced information relative to instructed information. Koehler (1995) argued that the way in which base rate information is learned affects the way decision-makers use this information; specifically, when base rates are directly experienced through trial-by-trial outcome feedback, their impact on judgments increases. Spellman, (1993; 1996) also found that when base rates are learned in an experiential manner subjects show better base-rate use. Gigerenzer, Hell and Blank (1988) argued that a reason for observed variations in use of base-rate information in probabilistic judgment is the different influence that varying types of information have. In their studies using Kahneman and Tversky's (1973) famous engineer-lawyer problem Gigerenzer et al. replicated the classic base-rate neglect finding with a verbal presentation where base-rates were asserted. However, when their subjects directly experienced the base-rates themselves (via sampling experience rather than being instructed) base-rate neglect disappeared and subjects clearly used the base-rates in a proper fashion.

Such results suggest that *instructing* subjects that all the cities they would be seeing were among the 30 largest in Germany and that, of these 30, the 9 cities with soccer teams are larger than the 21 other cities in 78% of all possible pairs may well not enable subjects to exploit this information effectively—particularly if it is somewhat at odds with experienced information. The same information might be utilised if it was learned through direct experience. Newell and (2004) also criticised Goldstein and Gigerenzer's (2002) study for not allowing respondents to learn about the soccer club information incrementally instead of relying on subjects unproven ability to integrate information about percentages provided at training into their test decisions.

In our experiments we assumed that there was no such problem with the two bits of conflicting information—the conflicting information is presumably well understood: even those with no knowledge of English soccer teams will appreciate that a team leading at half-time has an advantage over its opponent. And, although in our study we were unable to say anything precise about the diagnostic value of the half-time information—we offered our subjects no advice about this—advice is not necessary in order to test whether or not it had any influence on judgment, which it plainly did.

In Goldstein and Gigerenzer's (2002) study it is not clear that the diagnosticity of the conflicting information was sufficient to over-ride the recognition cue, even if subjects had integrated it into their decisions. Bröder and Eichler (2006) argue that, for the critical test items in Goldstein and Gigerenzer's (2002) study, where respondents had to choose between a city they recognised but knew didn't have a soccer team and a city they didn't recognise but might have a football team, the actual predictive success rate of the soccer cue was much lower than 78%. As there is only a 5 in 22 chance that the unrecognised city has a soccer team (they knew 9 of the 30 cities had a soccer team and knew the identities of four of the cities which did and four which did not) Bröder and Eichler calculated the predictive success rate of the soccer team cue is only 0.56 relative to the estimated validity of recognition of 0.80.[7] Arguably the predictive

[6]We thank Dan Goldstein for this suggestion

[7]Bröder and Eichler's (2006) "Predictive success rate" is not the same as cue validity but reflects the fact that, for the clear majority

success of the soccer team cue is even lower than 0.56. Subjects would be entitled to infer that the chance that an unrecognised city had a soccer team was less than 5 in 22 if they understood that the two cues are likely to be intercorrelated such that cities they did not recognise are less likely to have soccer teams than cities that they did recognise. As a result an inference that unrecognised cities are even less likely to have soccer teams would be justified.

Accordingly Goldstein and Gigerenzer's (2002) result that recognition appeared to be used in a non-compensatory fashion may not be due to any special status accorded to recognition: not because people didn't use other information but rather because, in this specific situation, recognition simply outweighed the soccer cue. With these cue values, decision strategies as diverse as multiple regression or recognition-based decision-making would lead to the same outcome.

In our experiments we found that subjects were more likely to forecast a team to win when additional information was consistent with recognition than when it was not. Although Goldstein and Gigerenzer found that 92% of decisions made with conflicting information were in accord with the recognition cue they did not report the proportion of decisions made when the two bits of information were *not* conflicting. Nevertheless, in a replication of Goldstein and Gigerenzer's experiment, Newell and Fernandez (2006) found that subjects were significantly more likely to choose the recognized city when it was also known to have a soccer team than when it was known not to have a soccer team.

The data we present here confirm the impressive power of decisions based on recognition. However, our experiments indicate that recognition based decisions can be influenced by other information. Although our experiments avoid criticisms made of other studies by Pachur et al., the half-time scores are "givens"—presented information rather than information drawn from memory—which deviates from the memory-based test setting emphasised by Gigerenzer and Goldstein (2011). Nonetheless, our consideration of the evidence leads us to doubt the basis for the suggestion that "... no other information can reverse the choice determined by recognition" (Goldstein & Gigerenzer, 2002, p. 82).

In more recent writings the notion that the recognition heuristic is invoked whenever recognition can discriminate has been qualified. Gigerenzer and Goldstein (2011 p. 105) write: "We do not assume that people follow the recognition heuristic unconditionally", arguing that two processes, recognition and evaluation guide the adaptive selection of the recognition heuristic. Gigerenzer and Goldstein (2011) devote a section of their article to the evaluation process and give some criteria for selection of the heuristic; nevertheless, Gigerenzer and Goldstein (2011) concede that how the evaluation process works is not yet well understood. Some significant questions arise about how the evaluation process operates (see Newell, 2011) and also render testing of the recognition heuristic more difficult: how can we know if the recognition heuristic was applied in any case? Are inferences inconsistent with the recognition heuristic evidence against it or merely evidence that it was not invoked?

Gigerenzer and Goldstein's (2011) recent statement that "[w]e personally have no doubts that recognition is sometimes dealt with in a compensatory way, especially when the ecology favors doing so" (p. 110) appears to undermine the significance of our findings—but only at the expense of some of the advertised virtues of the recognition heuristic, namely simplicity and testability. Goldstein and Gigerenzer's (2002, p. 88) claim that fast and frugal heuristics are better formulated than representativeness and availability (Tversky & Kahneman, 1974) because they "... allow one to make quantifiable and testable predictions, and avoid possible misunderstanding (or mystification) of the processes involved, even if they do sacrifice some of the allure of the unknown" looks less persuasive than it once did.

of critical pairs, the unrecognized city would actually have no soccer team (an expected 17 out of every 22 cases). Hence using the soccer team cue on these cases would, by definition because the cue does not discriminate them, produce correct inferences only 50% of the time. The success rate would be 78% only for those (5 out of every 22) cases where the unknown city had a football team.

References

Andersson, P., Edman, J. & Ekman, M. (2005). Predicting the World Cup 2002: Performance and confidence of experts and non-experts. *International Journal of Forecasting, 21*, 565–576

Ayton, P. & Önkal, D. (1997). *Forecasting football fixtures: Confidence and judged proportion correct.* Unpublished manuscript, Department of Psychology, The City University, London.

Beaman, C. P., Smith, P. T., Frosch, C. A., & McCloy, R. (2010). Less-is-more effects without the recognition heuristic. *Judgment and Decision Making, 5,* 258–271.

Bröder, A. and Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica, 121*, 275–284.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 104,* 650–669.

Gigerenzer, G. & Goldstein, D. G., (2011). The recognition heuristic: A decade of research. *Judgment and Decision Making, 6*, 100–121.

Gigerenzer, G., Hell, W. & Blank, H. (1988). Presentation and content: The use of base rates as a continuous

variable. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 513–525.

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

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G., Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 37–58). Oxford: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of Ecological Rationality: The Recognition Heuristic. *Psychological Review, 109*, 75–90.

Hoffrage, U. (1995). *The adequacy of subjective confidence judgments: Studies concerning the theory of probabilistic mental models.* Unpublished doctoral dissertation, University of Salzburg, Austria.

Hoffrage, U. (2011). Recognition judgments and the performance of the recognition heuristic depend on the size of the reference class *Judgment and Decision Making, 6*, 43–57.

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Katsikopoulos, K. V. (2010). The less-is-more effect: Predictions and tests. *Judgment and Decision Making, 5,* 244–257.

Koehler, D. J. (1996). A strength model of probability judgments for tournaments. *Organizational Behavior and Human Decision Processes, 66*, 16–21.

Koehler, J.J.J. (1995). The base-rate fallacy reconsidered - descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*, 1–55.

Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., & Gigerenzer, G. (2010). From recognition to decisions: Extending and testing recognition based models for multi-alternative inference. *Psychonomic Bulletin & Review, 17*, 287–309.

Newell, B. R. (2011). Recognising the Recognition Heuristic for what it is (and what it‘s not). *Judgment and Decision Making, 5*, 409–412.

Newell, B. R. & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19*, 333–346.

Newell B. R. & Shanks D. R. (2004). On the Role of Recognition in Decision Making *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 923–935.

Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition, 90*, B1-B9.

Pachur, T. & Biele, G. (2007). Forecasting from ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica, 125*, 99–116

Pachur, T., Bröder A., & Marewski J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making, 21*, 183–210.

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 983–1002.

Pohl, R. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19*, 251–271.

Reimer, T., & Katsikopoulos, K. V. (2004). The use of recognition in group decision-making. *Cognitive Science, 28*, 1009–1029.

Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 150–162.

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review, 112*, 610–628.

Serwe, S, & Frings, C. Who will win Wimbledon? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making, 19*, 321–332.

Smithson, M. (2010). When less is more in the recognition heuristic. *Judgment and Decision Making, 5,* 230–243.

Spellman, B. A. (1993). Implicit learning of base rates. *Psycholoquy 4(61)* base-rate.4.spellman.

Spellman B. A. (1996). The implicit use of base rates in experiential and ecologically valid tasks. *Behavioral and Brain Sciences, 19*, 38.

Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

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

Yates, J. F., McDaniel, L. S. & Brown, E. S. (1991). Probabilistic forecasts of stock prices and earnings: The hazards of nascent expertise. *Organizational Behavior and Human Decision Processes*, *40*, 60–79.