Women and taxis and dangerous judgments: Content sensitive use of base-rate information

Lívia Markóczy[1]
Cranfield School of Management
Cranfield University
Cranfield, Bedford
MK43 0AL
Phone +44 (0)1234 751 122 (ext 3757)
Fax +44 (0)1234 750 070
L.Markoczy@cranfield.ac.uk

Jeffrey Goldberg
Cranfield Computer Centre
Cranfield University
Cranfield, Bedford
MK43 0AL
J.Goldberg@cranfield.ac.uk

MDEMXG

[1]We would like to thank Cliff Bowman, Mark Jenkins and Julia Newton for assisting the experiment.

Abstract

This study shows that the use or non-use of base-rate information in probability estimation depends not just on the form of the problem, but also on the content about which the base-rate information is given. When the information is stated about characteristics of types of humans it is used and recalled better than if the information is about arbitrary things. We speculate that the content specificity is a consequence of the human mind's attentiveness to stereotypes.

Making predictions for the uncertain future is part of our daily lives. We assess whether it is going to rain before deciding to take an umbrella, and we expect that traffic will stop at the read light before deciding to cross the road at an intersection. People also frequently make probability assessments as part of their job in organizations, ranging from judging the suitability of a candidate for a given job to making assessments for the future behavior of their fellow colleagues, competitors or customers. While these judgments are often made routinely and sometimes effortlessly there is no guarantee that these are also made optimally.

According to Bayesian reasoning people can only make optimal probability assessments if they take into consideration both what they know about the specific entity they are considering ("individuating information") and also rely on what they generally know about the broader population of which the entity belongs to (ie, "base-rate" or "prior" information). Doctors, for example, should assess the likelihood of a breast cancer for a patient with a positive mammogram, by not only considering the reliability of the mammogram (individuating information) but also considering the known likelihood that a women in the age group developing a breast cancer (base-rate information). Kahneman and Tversky (1973) and others (e.g., Lyon and Slovic, 1976; Nisbett and Borgida, 1975; Taylor and Thompson, 1982) found that people tend to deviate from the normative Bayesian model by ignoring base-rate information in the presence of even flimsy individuating information.

The claim that people generally ignore base-rate information is controversial largely because other studies have shown that at least under some circumstances people do take into account base-rate information (e.g., Carroll and Siegler, 1977; Manis, Dovalina, Avis and Cardoze, 1980; Bar-Hillel and Fischhoff, 1981). What is uncontroversial is that under some conditions people do not make proper use of base-rate information and at other times they do. As Fiske and Taylor (1991, p: 361) note "the appropriate question regarding the use of base-rate information is *when* is it used, rather than if it is used." More ambitiously, we want more than just a catalog of when base-rate information is used. We should work toward a theory of human cognition that explains why it is used in some cases and not in others.

We have noticed that when the object about which the base-rate informa-

tion is given is about known "types" of human beings base-rate information is more likely to be used (e.g., Bar-Hillel and Fischhoff, 1981; Manis et al., 1980; Davis, 1979; Nelson, Biernat and Manis, 1990).[1] Here we report a test of whether base-rate information is more likely to be used when the object of the base-rate information is a type of human instead of some non-human object. Our study described below confirms our view.

We defer most of the discussion for why we expected this result to the second part of the paper. A brief overview of our rather speculative view is that there is more flexibility for non-representativeness for humans than for non-humans. Paradoxically, it is because we are more willing to take an individual human as non-representative of their class that increases the force of stereotypes for humans. With an inanimate object, may treat the degree to which it is representative of its class as an all or nothing thing. Once we have individuating information our inflexibility with regard to its representativeness leads us to reject the base rate information altogether. With types of humans we are more flexible. We will accept a certain degree of variation and "individualness" without it conflicting with what we believe about certain kinds of people.

It should be noted that we are not claiming that human versus non-human objects account for all of the mixed findings in the base-rate literature. Base-rate is also used when it is presented to the subjects as frequencies (e.g., Kahneman and Tversky, 1979; Gigerenzer, 1996; Gigerenzer and Hoffrage, 1995; Cosmides and Tooby, 1996; Kahneman and Tversky, 1996), when the casual relevance of the base-rate information is made salient (Ajzen, 1977; Tversky and Kahneman, 1980), or when the problems are framed as repetitive rather than as unique (Kahneman and Tversky, 1979).[2]

[1] None of those studies particularly attributed the use of base rate information to being about people, but we feel that their results can be better understood in these terms.

[2] It should be noted that at least some of the work demonstrating that people are good with base rates when presented with frequency information is open to other interpretations. Cosmides and Tooby's (1996) criticize much of the work on the base-rate fallacy by Tversky and Kahneman, pointing out that many of failures to take into account base rate information may be due to other facts about the experiments. Howson and Urbach (1993, p. 421–422) commenting on (an earlier version of) Cosmides and Tooby (1996) point out "...if caution is advised over Kahneman and Tversky's interpretation of their results, then

The main study

We used a 2×3 factorial design, in which subjects were either given a problem with a low base-rate of 15 percent or a high base-rate of 85 percent. The individuating information (70 percent) was held constant.[3] The three factors were whether the base-rate information was given about colors of taxicabs, success of women, or future honesty of people with criminal records. An example of the the high/taxi case is listed in Figure 1, and Figure 2 contains the variable part of the texts for all of the treatments. The taxicab case is a variant of the well known case by Kahneman and Tversky (1973) while the other cases were constructed to be analogous with the taxicab case.

In what follows we will refer to the three types of objects listed for base-rate information as "taxis", "women", "criminals" and the two base-rates of 15% and 85% as "low" and "high". Collectively women and criminals referred to as "humans".

We had two initial predictions. The primary one was that base-rate information would be taken into more account by more people when it is given about humans than about taxis. This prediction is fully supported by our results. We were concerned in the design of the study that social desirability effects would mask some of the primary effect with regard to women. We were concerned that out of a desire to not appear sexist the overall scores about women would be higher for both the low and high case than when about criminals. It turns out that we needn't had worried. The difference in results for the women and the criminals was not very large and to the extent that it existed was mostly opposite as expected.

There were a few other things, which although not predictions, we were not surprised to find. The results show up marginally more clearly for the

it is positively urged for Cosmides and Tooby's interpretation of theirs, which is no less than that people do possess a faculty for probabilistic, even Bayesian, reasoning, which is manifested however for a purely frequentist concept of probability. A more accurate conclusion is that their respondents are competent at whole number arithmetic... But with *probabilistic* reasoning... Cosmides and Tooby's results have very little to do at all, despite their dramatic claims."

[3] To keep the description of the information simple, we stated the individuating information as shown and did not separate positive and negative hit rate.

Thank you for participating in this study on judgment under uncertainty. This is not a test, and is completely anonymous. It should take a few minutes at most. When you are done, please fold this sheet in half. Please work alone and be aware that not everybody has been given the same question. Please do not discuss this with people from other classes who have yet to participate. Once the study is complete, you will be told the results. Again, thank you for your participation.

You are an investigator checking out a crime report, and at this point you are trying to determine the colour of the taxicab a suspect fled in. You know that 85 percent of the taxicabs in your city are green.

A witness reported seeing the suspect flee in a green taxicab at dusk; you know that witnesses report the color correctly 70 percent of the time for similar conditions, and incorrectly 30 percent of the time.

Please give an estimate of the probability that the colour of the cab the suspect fled in was green. Indicate your estimate by placing an "×" on the scale below.

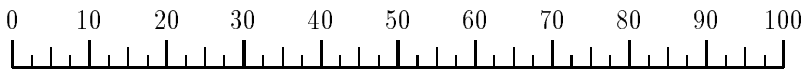

Figure 1: One of the six possible questions given to subjects

cab You are an investigator checking out a crime report, and at this point you are trying to determine the colour of the taxicab a suspect fled in. You know that [15 or 85] percent of the taxicabs in your city are green.

A witness reported seeing the suspect flee in a green taxicab at dusk; you know that witnesses report the color correctly 70 percent of the time for similar conditions, and incorrectly 30 percent of the time.

Please give an estimate of the probability that the colour of the cab the suspect fled in was green.

women You are helping to review job applications for a company, and you are trying to determine whether a particular female job candidate will be successful in the job. You know that [15 or 85] percent of women succeed in this kind of job.

The candidate passed a test which may predict success in the particular job; you know that the test predicts correctly in 70 percent of the cases, and incorrectly in 30 percent of cases.

Please give an estimate of the probability the candidate will succeed in the job if hired.

convicts You are helping to review job applications for a company, and you are trying to determine whether a particular candidate with a criminal record will act honestly in the job. You know that [15 or 85] percent of people with criminal records act honestly in this kind of job.

The candidate passed a test which may predict honesty in the particular job; you know that the test predicts correctly in 70 percent of the cases, and incorrectly in 30 percent of cases.

Please give an estimate of the probability the candidate will act honestly if hired.

Figure 2: Variable part of texts for the three (six) varieties of questions

women than for the convicts. We do believe that the effect should be stronger when the base-rate is given about what appear to be "natural kinds" of humans than more accidental kinds. We discuss this more in the concluding remarks. While we are not surprised by this tendency, we had thought that it would have been compensated for by political desire of the subjects to consciously use only the individuating information for women. Finally, we are again not surprised that the influence of base-rate was stronger for the low base-rates than for the high ones. Fischhoff, Slovic and Lichtenstein (1979) found similar results and we discuss why that might be the case for this study later.

Pilot

Prior to running the experiment on our primary sample, we ran smaller tests on two groups of 57 subjects in total, but slightly different conditions for each group. Subjects included students from an executive-MBA course from the same university as for the main study. Our results in the pilots were all in the same direction as the results in the main study, but were noisier. We modified two things about the procedures to reduce some of this noise.

It turns out that in some pilot groups, individuals took more than 10 minutes working on the problem, apparently attempting to calculate an answer. This led to a large number of answers of 60 for the high cases and 10 for the low cases. (These are approximately 70% of 85 and 70% of 15 respectively.) As a consequence, we added to the verbal instructions that we don't want people to calculate answers, but to give their impressions. This did help reducing the numbers of 60s and 10s in the final study, but many still remain. See Figures 3 and 4 for histograms of the main data.

It also appears that where the administrator of the experiment (the instructor in the class where it was carried out) was female the results for the women case was even noisier (higher variance), although we can't say that with much confidence. Still, we arranged to have a male administrator and to conceal gender information about designers of the study.

Additionally, there were some small changes to the wording about marking scales and correction of typographical errors after some of the pilots. If

Base-rate	Taxi	Women	Criminals	Total
High	37	28	21	86
Low	37	25	24	86
Both	74	53	45	172

Table 1: N for different cases

we agglomerate the data for the three pilots we find – in the essential features – the same results that we have for the main study. Because of the slight changes in written and verbal instructions, however, we find it is simpler to report on the main study alone.

The Sample

The sample consists of 172 MBA students at a major UK business school. All have at least 3 years of business experience. Their average age is 31 years, ranging from 25 to 47. The experiment was administered within the first 3 weeks of the first term, so the students had not been exposed to Bayesian analysis or anything similar as part of their program at that point.

In anticipation of agglomerating the responses for the "women" and "criminals" treatments we worked to ensure that there were more "taxi" cases than others. A clerical error in the production and shuffling of the response sheets lead to more "women" cases being given out than "criminal" cases. These numbers are summarized in table 1.

We deliberately did not collect any information about the respondents. This was both to ensure them that their responses were entirely anonymous, as well as not encouraging them to think that the study was about sexism or attitudes towards criminals.

Results and Analysis

Our prediction is that for the human cases (criminals and women) we should find that base-rate information plays a more important role. Inspection of table 2 bears this out. The difference between the mean score for the low

	Taxi		Women		Criminals	
Base-rate	**Mean**	**s.d.**	**Mean**	**s.d.**	**Mean**	**s.d.**
High	63.69	7.94	78.21	12.37	72.33	13.45
Low	43.91	25.08	24.00	19.07	28.98	24.04
Both	53.45	21.24	52.64	31.38	49.21	29.33
High−Low	20.50		54.21		43.35	

Table 2: Descriptive statistics of scores

	Taxi			Human		
Base-rate	N	**Mean**	**s.d.**	N	**Mean**	**s.d.**
High	37	63.69	7.94	49	75.69	13.17
Low	37	43.91	25.08	49	26.44	21.79
Both	74	53.45	21.24	98	51.07	30.51
High−Low		20.50			49.26	

Table 3: Scores for combined cases

base-rate in the taxi case is 20.50 points lower than for the high base-rate case, showing that the base-rate information was taken into some account in some cases. For the criminal case, that difference is more than twice as large (43.35) and larger still for the women case (54.21). Even with large variance of scores in each condition (see table 2), these results are highly significant.

In the analysis that follows we combine the criminal and woman cases since they both show the expected effect, although the effect shows up more clearly in for the women case than the criminal case. This agglomeration of the human cases does not change the overall results, but makes the analysis and reporting easier.[4] Table 3 provides a description of the agglomerated results.

With the combined data, we can think of the study as a 2×2 factorial design, which can be subject to analysis by simple ANOVA. Table 4 shows via the interaction effect that the effect of base-rate differences is greatly

[4]The full data set is available from the authors. for those who wish to conduct analyses not reported here.

	Mean square	df	F	p
High/Low	51288.655	1	150.161	< 0.0005
Human/Taxi	237.397	1	.695	.406
Interaction	8715.638	1	25.517	< 0.0005
Explained	22483.539	3	65.827	< 0.0005
Residual	341.557	168		
Total	730.013	171		

Table 4: ANOVA for interaction between base-rate and type of object (human/non-human)

enhanced in those cases where the base-rate information was given about humans ($F = 25.517$, $df = 1$, $p < .0005$).

Better answers?

While it is the case that those given problems where the base-rate information was about humans took that base-rate information more into account than others, it still remains a question as to whether this improved their judgment. The correct (Bayesian) judgments using all and only the information specified in the problems is 29.17 for the low base-rate and 92.97 for the high base-rate.[5] All things being equal and given the results reported above, we should expect those with the human objects would judge more accurately than those judging about cabs. This is broadly true, but not as strong as one might otherwise expect. Table 5 is analogous to table 3 except that it reports on

[5]There are some, most notably Gigerenzer (1996) and Cosmides and Tooby (1996), who object to talking about "correct" and "incorrect" responses, or of "biases" and "errors" in human reasoning. If you share their objection then read "normatively defined by standard Western statistical analysis" for "correct" and "natural tendency to optimize some forms of reasoning for particular situations" for "biases". But rewording doesn't change the facts. And since there are a vast number of ways to calculate something correctly, the best way to test hypotheses about the actual architecture of the mind is to understand when and where errors occur. That is, we need to see what makes a task easy for the human mind and what makes a task difficult in order to build up a clear theory of the mind's design.

	Taxi			Human		
Base-rate	N	**Mean**	**s.d.**	N	**Mean**	**s.d.**
High	37	29.28	7.94	49	17.69	12.61
Low	37	24.78	14.55	49	20.29	8.41
Both	74	27.03	11.93	98	18.99	10.80

Table 5: Distance from correct answers

absolute distances from the Bayesian answers.

By inspection it is clear that as predicted judgments were more accurate in the human cases than the taxi ones ($t = 4.593$, $df = 170$, $p = .00004$), however this effect was far more pronounced for the high base-rate case ($t = 4.847$, $df = 84$, $p = .000003$) than for the low base-rate case ($t = 1.778$, $df = 84$, $p = .040$). It appears that use of the base-rate information in the high case pushed the responses closer to, but not over, the correct response (5 out of 49 of the high/human respondents were above 92); while in the human low base-rate case many (32 out of 49) of the responses fell below the correct answer.

The low taxi respondents were also closer to the correct answer than the high cab respondents as can be seen from table 5 ($t = 1.63$, $df = 72$, $p = .054$ (one-tailed)). This result is consistent with previous observations that low base-rate information is more considered than high base-rate information (Fischhoff et al., 1979), however this can also be explained more mundanely by noting that a number of respondents selected 10 for low base-rates and 60 for high base-rates (simply multiplying the base-rate with the reliability of the individuating information). For the high base-rate case this shifted an average near the individuating reliability (70%) downward, *away* from the correct response. While for the the low base-rate, it also shifted the result downward, *toward* the correct answer.

How is content used?

We have deferred all discussion of why the nature of the object about which the base-rate information should matter, and specifically about why it should

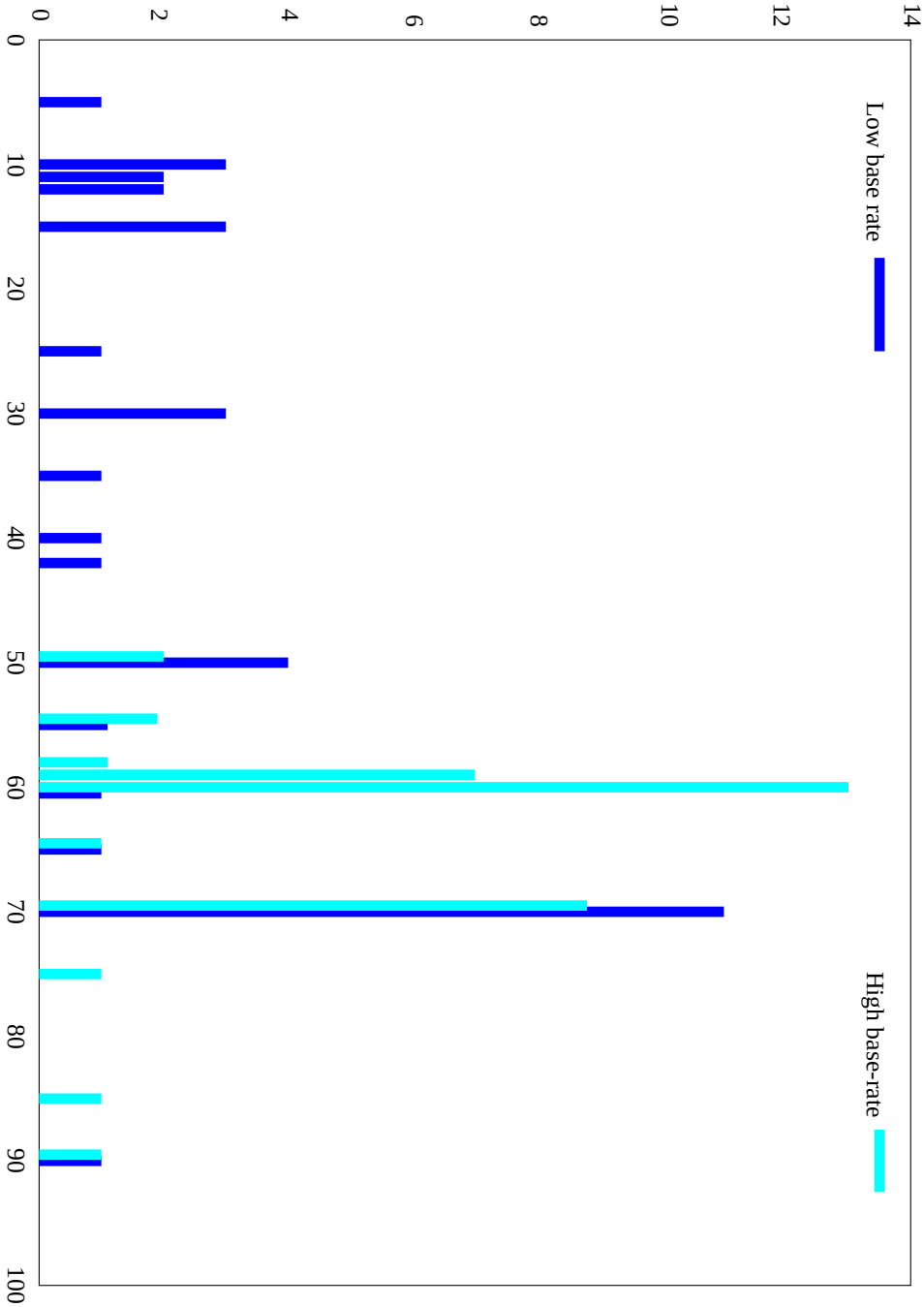

Figure 3: Histogram for the high and low cab cases

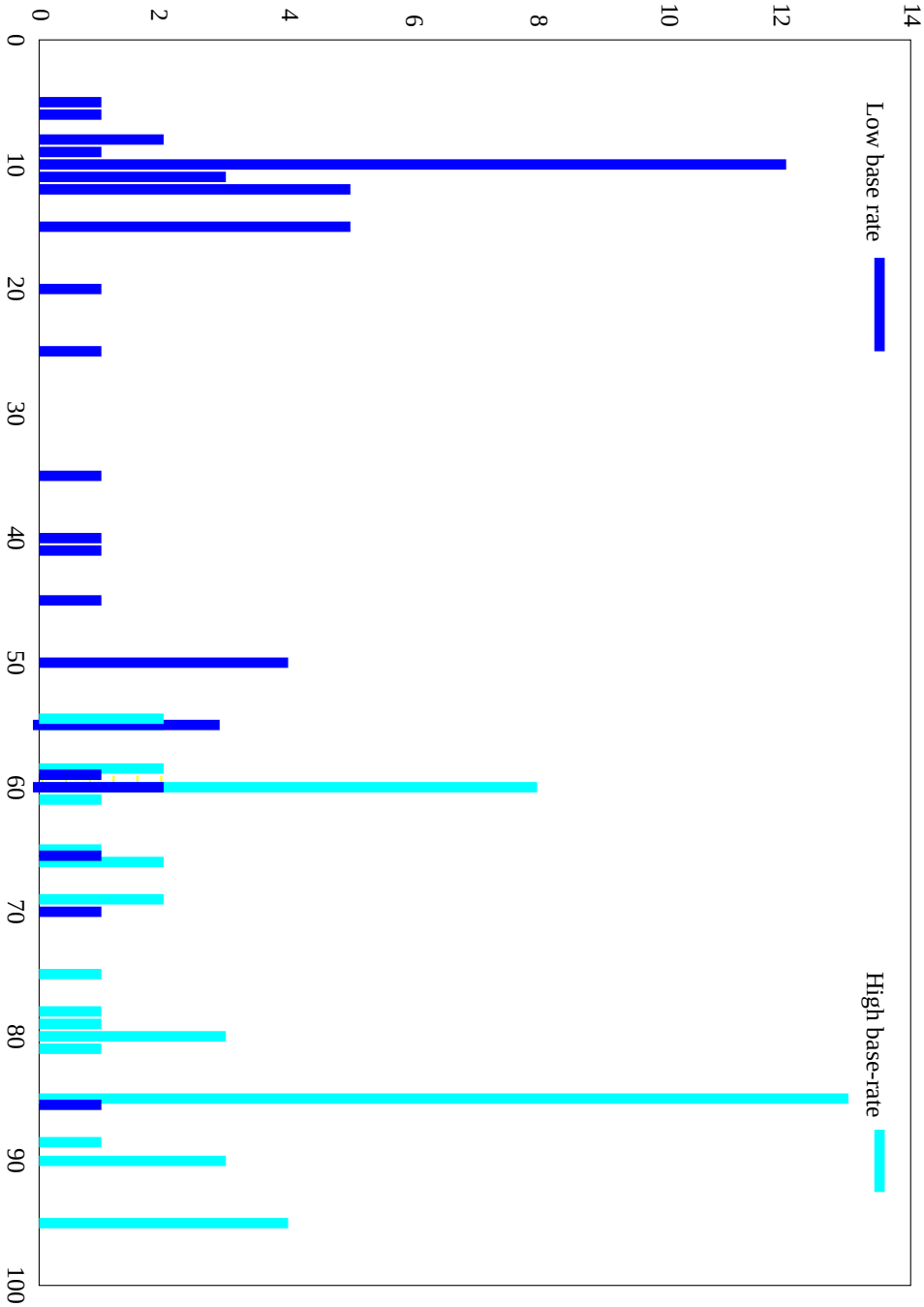

Figure 4: Histogram for the high and low human cases

matter in the particular way it does. We were confident enough that it would matter this way to design a specific study to test it, but our theoretical basis is somewhat speculative.

The over all scheme is that parts of reasoning mechanisms are not independent of content. In stating why we expect this sort of content sensitivity, it is best to do so in comparison to other cognitive tasks that have shown content sensitivity. In particular Wason's selection task (Wason, 1966) has been shown to exhibit very strong content sensitivity. Consider the task in Figure 5. If you are unfamiliar with this task, please try to answer it before reading further. Wason and others (e.g., Wason, 1966; Wason and Johnson-Laird, 1972) discovered that people are very bad at this task. Most subjects incorrectly selected either the "D" card alone or the "D" and "3" cards. However, when given the logically equivalent task stated in Figure 6, subjects do much better. The content specificity had been long noted and is discussed by Johnson-Laird (1983, p. 29–34) to argue that reasoning does not closely parallel the symbol manipulation of simple logic. He suggested that content affects reasoning through the ease with which people construct a working model of the problem in their minds. These so called *mental models* are the ways human beings understand the world around them. Individuals seem to find it less difficult to construct valid mental models if they are familiar with the problem (Johnson-Laird, 1983), if the problem involves application of a social rule (Cosmides and Tooby, 1992) or if the information is presented in frequency formats (Gigerenzer and Hoffrage, 1995).

There are two ways for content specificity to occur: (1) the base-rate information is incorporated or not into a mental representation depending on content; or (2) the base-rate information is incorporated in both cases, but only used or not depending on content. We suspect that the mind would not be designed to incorporate information it won't use; therefore the information "that 85% of the taxicabs in a city are green" is less likely to be incorporated into one's mental model of the city than the information "that 85% of women will succeed in a particular job" is to be incorporated into a model of women job candidates. We ran a small variant of the main study to test this.

Experiment 2

The basic idea of this variant is that if base-rate information is differentially incorporated into a mental model depending on content, then it will be remembered differently.

Twelve management research students (Ph.D and MPhil) participated in this experiment from the same university where the main experiment was conducted. Both phases were conducted in a class-room setting.

The experiment was done in two stages. The first stage was almost identical to the main experiment. The only differences were that only two of the six possible cases were distributed (high/taxi and high/criminal) and that the gender of the designer of the experiment was not concealed from the subjects.[6] Results of phase one were in the same direction as in the main study but not significant, which was to be expected given the sample size and design. Subjects were given no indication until the second phase that there would be a second phase.

The second phase of the experiment was conducted 14 days after the first phase. Subjects were asked to indicate, by a show of hands, whether they had the problem involving taxis or criminals. They reported no difficulty in remembering which question they had. They were then given a copy of the same sheet that they had for phase one, except that the base-rate information, 85%, was obscured. They were given verbal instructions to fill in this blank space based on what they remembered from the original problem. The results are summarized in Table 6. Those who had the criminal case recalled the base-rate information more accurately (deviated less from the correct answer of 85%) than those who had the taxi information ($t = 2.079$, $df = 10$, $p = .032$ one-tailed).

While we have not ruled out many alternative explanations, this result supports our view that people are more likely to correctly incorporate base-rate information into their mental models when that information is about types of people than when it is about non-humans.

[6]This latter point should make no difference since the case about women was not used here.

	Taxi		Criminal	
	x	$\|x-85\|$	x	$\|x-85\|$
	80	5	85	0
	80	5	80	5
	80	5	85	0
	70	15	80	5
	70	15	80	5
	60	25		
	90	5		
Mean	75.71	10.71	82.00	3.00
s.d.	9.04	7.28	2.45	2.45

Table 6: Accuracy of recall of high base-rate for two problems

The context of content specificity

Johnson-Laird (1983, p. 33) states when discussing the content specificity of the Wason selection task that "Memory is plainly important for correct performance in the selection task: no effect of content can be explained without appeal to previous experience." He has overlooked an important class of potential explanations for content specificity. The alternative to memory based explanations is that we are innately content sensitive and, more particularly, that these innate structures are adaptations or the consequences of adaptations:[7]

> Because we know that the human mind is the product of the evolutionary process, we know something vitally illuminating: that aside from those properties acquired by chance, the mind consists of a set of adaptations.... Such a view is uncontroversial to most behavioral scientists when applied to such topics as vision or balance. Yet adaptionist approaches... are considered

[7] We don't wish to discuss or argue the distinction between an adaptation and the consequence of an adaption here. It is an informal distinction anyway. To ground the intuition: Iron in blood is an adaptation; redness of blood is the consequence of an adaptation.

Part of your new clerical job at the local high school is to make sure that student documents have been processed correctly. Your job is to make sure that the documents conform to the following rule:

"If a person has a "D" rating, then his document must be marked code "3".

You suspect the secretary you replaced did not categorize the students' documents correctly. The cards below have information about the documents of four people who are enrolled at this high school. Each card represents one person. One side of a card tells a person's letter rating and the other side of the card tells that person's number code.

Indicate only those card(s) you definitely need to turn over to see if the documents of any of these people violate the rule.

D	F	3	7

Figure 5: Example Wason selection task, from Cosmides and Tooby (1992)

> radical... when applied to most other areas of human thought and action.... Nevertheless, the logic of the adaptationist position... leads to the expectation that humans should have evolved a constellation of cognitive adaptations to social life. Our ancestors have been members of social groups and engaging in social interactions for millions and probably tens of millions of years. To behave adaptively, they not only needed to construct a spatial map of the objects disclosed to them by their retinas, but a social map of the persons, relationships, motives, interactions, emotions, and intentions that made up their social world. [Cosmides and Tooby, 1992, p. 163]

No-one is saying that we evolved to have rules about drinking ages in our heads, but instead that we have evolved to be able to integrate and comprehend rules about social structures. That is, we are especially attuned to identifying those who violate rules about social exchange, whatever those particular rules happen to be in a particular society or situation. Cosmides and Tooby (1992) show that people perform better on the Wason's selection

In a crackdown against drunk drivers, Massachusetts law enforcement officials are revoking liquour licenses left and right. You are a bouncer in a Boston bar, and you'll lose your job unless you enforce the following law:

"If a person is drinking beer, then he must be over 20 years old."

The cards below have information about four people sitting at the table in your bar. Each card represents one person. One side of the card tells what a person is drinking and the other side of the card tells that person's age.

Indicate only those card(s) you definitely need to turn over to see any of these people are breaking the law.

drinking beer	drinking cola	25 years old	16 years old

Figure 6: Example Wason selection task for social exchange, from Cosmides and Tooby (1992)

task when the rule that needs to be checked is some sort of social rule where violations mean a certain sort of cheating.[8]

If predicting how people will behave is connected to a different mental module or faculty than some general reasoning ability, we might expect that the sorts of results we found. Base-rate information about categories of people may naturally be used, along with individuating information or not as the case may be, in judging the future behavior of people.

It has been well demonstrated by studies on the use of stereotypes that consideration of base-rates improves when subjects are asked to make judg-

[8]The Cosmides and Tooby study has been widely reported and much discussed (e.g., Cronin, 1991; Ridley, 1996; Pinker, 1997; Liberman and Klar, 1996), and we suspect that many who are now familiar with the Wason selection task know of it exclusively through their work. But it must be noted that the content specificity of the task cannot all be explained in terms of social constraint. Johnson-Laird (1983) discusses a study (Wason and Shapiro, 1971) which also used rules like "When I travel [from London] to Manchester, I travel by train", and presented subjects with cards that had destinations on one side (Manchester or Leeds) and mode of travel on the other (train or car). They found that with this rule subjects did much better. Although Cosmides and Tooby (1992) have come under sharp criticism (e.g., Liberman and Klar, 1996), we believe that they are ultimately on the right track and use a similar sort of explanation for our content specificity.

ments about humans. Nelson et al. (1990), for example, found that subjects judge the heights of males and females differently based on their stereotypical concepts of height differences between the two sexes. From this study, however, we could not conclude that stereotypical information is only considered when judgments are made about people as they failed to run a control where subjects would have been asked to judge the heights of non-human objects. More problematic for our argument are those studies which did not find that stereotypic information is used. Most notably among these is reported by Locksley, Borgida, Brekke and Hepburn (1982) and Locksley, Hepburn and Ortiz (1982) who found that although people believe that men are more likely to be aggressive than women, when presented with some individuating information judgments of whether the person is aggressive or not did not make use of the base-rate information. Rasisnki, Crocker and Hastie (1985), however, demonstrated that these studies suffer from serious flaws. The most important of the criticisms is that when someone judges whether a person is aggressive they very well may be reporting that a particular person is "aggressive for a woman" or "not very aggressive for a man". If, for example, you are asked to judge whether a particular hippopotamus is graceful you will probably use different standards of gracefulness than if you are asked the same question about a horse.

Flexible representativeness

Another way to look at results is to suggest an initially paradoxical proposal: We are more likely to make use of base-rate information about types of people because we expect greater variation within each type. We may for instance consider a taxi to be representative of its class or not. It's representativeness is an all or nothing thing. With people, however, we may be able to tolerate some deviation from the norm for some particular individual without having to abandon the norm all together.

In this way, it is the expectation of variance in how representative types of people are of their type that allows us to still make use of information about the type when we have some individuating information.

Limitations

No one study can cover all bases and control for everything. As the discussion above suggests, even if we are completely correct about the interpretation of our results there is a great deal of other work that will need to be done to clarify more of the mechanism and determine exactly what features of content are relevant. Additionally we need to develop a strong theory of why content matters in the particular way that it does instead of merely providing a plausibility argument. We need to work to make the ideas above precise and test and refine further.

But are there reasons to suspect that even this is not modest enough? Could there be an entirely different account of our result? Possibly. Some studies on base-rate have shown that base-rate information has a greater effect on judgments where there is seen to be a plausible causal link between the category of the base-rate information and the ultimate outcome (e.g., Ajzen, 1977; Bar-Hillel, 1980). Subjects may very well see the plausibility of a causal link between being a criminal and being honest, or between being a women and being successful, but there is no causal link between being a taxi and being green. We do not feel that this poses as large a problem as it initially seems to. In many cities taxis are regulated and there in a clear link between being a taxi and being a certain color. No one near London or New York should be unfamiliar with the link between taxihood and color (black in one and yellow in the other). Basically there is nothing implausible about a causal link between color and being a taxi. One might counter that the link between commercial status of a vehicle and its color is not a *natural* link, while the relationship between being a convict and dishonesty is natural. That, however, is exactly our argument. The human mind takes some categories as naturally relevant to behavior, and so uses that category information. Because links between types of people and how they behave are seen by the mind as natural, base-rate information about types of people is used.

There are number of studies which at first glance appear to be counter examples to our claim that consideration of base-rates improves when judgments is made about humans. These studies tend to rely on variations of

the well known study by Kahneman and Tversky (1973) where base-rate information on population contained information on humans, say, 15 percent engineers and 85 percent lawyers. In these studies base-rate information was not always used in determining whether some particular individual is an engineer or a lawyer (e.g., Argote, Devadas and Melone, 1990). What sets these studies apart from our study is that for us the base-rate information was about some characteristic of a known type of person. While in these other studies the base-rate information was about which types existed in an arbitrary population. Our population was already a specific known type of individual and the base-rate information was about what portion of those individuals possess a particular characteristic. It is precisely this sort of base-rate information that we expect (and found) people to be good at using.

Suggestions for future studies

Once we recognize the importance of content specificity we should ask whether its importance is limited to judgments about humans or that exactly what aspects of content matter. That question must be the subject of future studies, but some particular suggestions of things to look at are listed.

1. *Factoring out cause.* In order to isolate the sense that there is a causal relationship between the types of people and their behavior, it would be worthwhile to replicate this study with a more arbitrary relationship between the population about which the base-rate information and the characteristic that is to be judged. For example subjects could be told that women have each been given cards that have the digit "1" on one side in 15 or 85 percent of the cases. And that these cards have the digit "2" in the remaining cases. Furthermore, they are told that a card with an "A" on one side predicts a "1" on the other side correctly in 70 percent of the cases, etc. Another group of subjects could be told that some uninteresting object, taxicabs for example, have been labelled with cards in this fashion. The purpose of this would be to test whether there is a difference in the correct use of the base-rate

information, even when the causal connections between the population and the characteristic under consideration are identical.[9]

2. *Human vs. animate.* We have suggested that when base-rate information is given about behavior of types of people the information is used. However, it may not be limited to people, but may extend to anything which "behaves". For example, if the problem were stated about the likelihood of chihuahuas being trainable to not bark at mail carriers, would the results have been like the cab case or the women case? Would the results be different for wild animals than for pets which are parasitic off of our desires to bound (Archer, 1997)?

3. *Natural kind vs. accidental kind.* Would we get different results if the base rate information were stated about people who have an odd number of consonants in their family names? That is, where the type of person about which the information is given is an arbitrary category, as opposed to a kind of category that the mind automatically recognizes as a category. If so, what are these categories: Perceived race, nationality, sex, age, eye color, handedness?

4. *Under use of individuating information about non-humans.* If we are correct that representativeness for the non-human cases is "an all or nothing thing" and the formulation of the problem forced the "nothing" alternative, then it should be possible to construct an experiment in which for non-humans the base rate information is used and the individuating information is ignored, while for the humans kinds the to sources of information are integrated.

Concluding remarks

Our main study suggests that the use or non-use of base-rate information is content sensitive. Our second study indicates that content sensitivity matters because it affects the ease with which subjects are able to construct the

[9]The actual design of this experiment is non-trivial. As stated here it is plausible that subjects might see the population not as women and taxis, but merely as cards.

correct mental model of a problem. We believe that the particular content sensitivity is the consequence of the evolved nature of the human mind. We evolved to making judgments about behaviors of types of people and so when information is given about types of people we incorporate that information into our mental models.

Coming to a better understanding of the nature of human decision making should eventually enable us to anticipate and correct for our natural biases in these judgments, thus avoiding the sorts of costly errors discussed by Slovic (1972).

Obviously our research is at an early stage, and much more needs to be done. But these first results are overwhelming and we look forward to a very interesting series of studies over the next few years.

References

Ajzen, I. (1977). 'Intuitive theories of events and the effects of baserate information on prediction'. ***Journal of Personality and Social Psychology***, **35**, 5, 303–314.

Archer, J. (1997). 'Why do people love their pets?'. ***Evolution and Human Behavior***, **18**, 4, 237–259.

Argote, L., Devadas, R. and Melone, N. (1990). 'The base-rate fallacy: Contrasting processes and outcomes of group and individual judgment'. ***Organizational Behavior and Human Decision Processes***, **46**, 296–310.

Bar-Hillel, M. (1980). 'The base-rate fallacy in probability judgments'. ***Acta Psychologica***, **44**, 211–233.

Bar-Hillel, M. and Fischhoff, B. (1981). 'When do base rates affect predictions?'. ***Journal of Personality and Social Psychology***, **41**, 4, 671–680.

Carroll, J. S. and Siegler, R. S. (1977). 'Strategies for the use of base-rate information'. ***Organizational Behavior and Human Decision Processes***, **19**, 392–402.

Cosmides, L. and Tooby, J. (1992). 'Cognitive adaptations for social exchange'. In Barkow, J. H., Cosmides, L. and Tooby, J. (Eds), ***The adapted mind***. Oxford: Oxford University Press. chapter 3, pp. 163–228.

Cosmides, L. and Tooby, J. (1996). 'Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty'. ***Cognition***, **58**, 1–73.

Cronin, H. (1991). ***The Ant and the Peacock: Altruism and sexual selection from Darwin to today***. Cambridge: Cambridge University Press.

Davis, D. A. (1979). 'What is in a name? a Baysian rethinking of attributional biases in clinical judgment'. ***Journal of Consulting and Clinical Psychology***, **47**, 1109–1114.

Fischhoff, B., Slovic, P. and Lichtenstein, S. (1979). 'Subjective sensitivity analysis'. ***Organizational Behavior and Human Decision Processes***, **23**, 339–359.

Fiske, S. T. and Taylor, S. E. (1991). ***Social Cognition***. McGraw-Hill Series in Social Psychology. second edn. New York: McGraw-Hill.

Gigerenzer, G. (1996). 'On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996)'. ***Psychological Review***, **103**, 3, 592–596.

Gigerenzer, G. and Hoffrage, U. (1995). 'How to improve Bayesian resoning without instruction: Frequency format'. ***Psychological Review***, **102**, 4, 684–704.

Howson, C. and Urbach, P. (1993). ***Scientific Reasoning: The Baysian approach***. Second edn. Chicago: Open Court.

Johnson-Laird, P. N. (1983). ***Mental models***. Cambridge, England: Cambridge University Press.

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

Kahneman, D. and Tversky, A. (1979). 'Intuitive predictions: Biases and corrective procedures'. ***TIMS Studies in Management Science***, **12**, 313–327.

Kahneman, D. and Tversky, A. (1996). 'On the reality of cognitive illusions'. ***Psychological Review***, **103**, 3, 582–591.

Liberman, N. and Klar, Y. (1996). 'Hypothesis testing in Wason's selection task: social exchange cheating detection or task understanding'. ***Cognition***, **58**, 127–156.

Locksley, A., Borgida, E., Brekke, N. and Hepburn, C. (1982). 'Sex stereotypes and social judgment'. ***Journal of Personality and Social Psychology***, **39**, 5, 821–831.

Locksley, A., Hepburn, C. and Ortiz, V. (1982). 'Social stereotypes and judgments of individuals: An instance of baserate fallacy'. ***Journal of Experimental Social Psychology***, **18**, 23–42.

Lyon, D. and Slovic, P. (1976). 'Dominance of accuracy information and neglect of base rates in probability estimation'. ***Acta Psychologica***, **40**, 287–298.

Manis, M., Dovalina, I., Avis, N. E. and Cardoze, S. (1980). 'Base rates can affect individual predictions'. ***Journal of Personality and Social Psychology***, **35**, 2, 231–245.

Nelson, T. E., Biernat, M. R. and Manis, M. (1990). 'Everyday base rates (sex stereotypes): Potent and resilient'. ***Journal of Applied and Social Psychology***, **59**, 4, 664–675.

Nisbett, R. E. and Borgida, E. (1975). 'Attribution and the psychology of prediction'. ***Journal of Personality and Social Psychology***, **32**, 5, 932–943.

Pinker, S. (1997). ***How the mind works***. New York: W. W. Norton.

Rasisnki, K. A., Crocker, J. and Hastie, R. (1985). 'Another look at sex stereotypes and social judgments: An analysis of the social perceiver's use of subjective probabilities'. ***Journal of Personality and Social Psychology***, **49**, 2, 317–326.

Ridley, M. (1996). ***The Origins of Virtue***. London: Viking, Penquin Books.

Slovic, P. (1972). 'Psychological study of human judgment: implications for investment decision making'. ***Journal of Finance***, **27**, 4, 779–799.

Taylor, S. E. and Thompson, S. C. (1982). 'Stalking the elusive "vividness" effect'. ***Psychological Review***, **89**, 2, 155–181.

Tversky, A. and Kahneman, D. (1980). 'Causal schemata in judgment under uncertainty'. In Fishbein, M. (Ed), ***Progress of social psychology***. Hillsdale, NJ: Erlbaum. pp. 49–72.

Wason, P. (1966). 'Reasoning'. In Foss, B. M. (Ed), ***New horizons in psychology***. London: Penguin.

Wason, P. and Johnson-Laird, P. N. (1972). ***Psychology of reasoning: Structure and content***. Cambridge, Mass.: Harvard University Press.

Wason, P. and Shapiro, D. (1971). 'Natural and contrived experience in a reasoning problem'. ***Quarterly Journal of Experimental Psychology***, **23**, 63–71.