Title
Why do people fail to consider alternative hypotheses in judgments under uncertainty?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors
Hayes, Brett K
Hawkins, Guy E
Newell, Ben R

Publication Date
2015

Peer reviewed

Why do people fail to consider alternative hypotheses in judgments under uncertainty?

Brett K. Hayes (b.hayes@unsw.edu.au)
School of Psychology, University of New South Wales
Sydney, 2052, Australia

Guy E. Hawkins (guy.e.hawkins@gmail.com)
Amsterdam Brain and Cognition Center, University of Amsterdam
Nieuwe Achtergracht 129, Amsterdam 1018WS, The Netherlands

Ben R. Newell (ben.newell@unsw.edu.au)
School of Psychology, University of New South Wales
Sydney, 2052, Australia

Abstract

Two experiments examined theoretical accounts of why people fail to consider alternative hypotheses in judgments under uncertainty. Experiment 1 found that a majority of participants failed to spontaneously search for information about an alternative hypothesis, even when this required minimal effort. This bias was reduced when a specific alternative was mentioned before search. Experiment 2 showed that when participants were given the likelihoods of the data given a focal hypothesis *p*(D|H) and an alternative hypothesis *p*(D|¬H), they gave estimates of *p*(H|D) that were consistent with Bayesian principles. The results show that neglect of the alternative hypothesis typically occurs at the initial stage of problem representation. However judgments are more consistent with Bayesian norms when they involve utilizing information about a given alternative.

Keywords: Judgment under uncertainty; Bayesian inference; Intuitive probability

Introduction

Judgments under uncertainty often involve evaluating the probability of a hypothesis given data. For example, when a person notices a persistent red blemish on their forearm they may want to judge the likelihood of skin cancer. Normatively, such judgments should involve consideration of the likelihood that the observed data could have arisen from a different cause. In its most elementary form, Bayes' theorem assumes that after observing a datum D (e.g., the red blemish), the probability of a focal hypothesis H (e.g., that the person has skin cancer), *p*(H|D), is assessed relative to the likelihood of the datum in the presence of the focal hypothesis *p*(D|H) *and* the likelihood of the data given alternative hypotheses *p*(D|¬H) (e.g., that the blemish is an allergic reaction).

A large body of evidence however suggests that people underweight or ignore alternative causes of the data in judgments under uncertainty (e.g., Beyth-Marom & Fischhoff, 1983; Dougherty, Thomas & Lange, 2010; Doherty, Mynatt, Tweney & Schiavo, 1979). Krynski and Tenenbaum (2007) for example, presented a version of the classic mammogram problem (cf. Eddy, 1982) where participants were told that a woman has received a positive mammogram and asked to estimate the (posterior) probability that she has cancer. To assist with problem solution participants were told the base rate of the focal hypothesis of having cancer, *p*(H), the likelihood of receiving a positive mammogram given cancer, *p*(D|H), and the false positive rate, or likelihood of a positive mammogram in the absence of cancer, *p*(D|¬H). When no additional information about false positives was supplied, less than 15% of participants generated a normative estimate of *p*(H|D) (cf. Hayes, Hawkins, Pasqualino, Newell & Rehder, 2014).

Such findings show that people often have difficulty in probabilistic reasoning involving alternative causes. However they reveal little about the mechanisms that give rise to these problems. A failure to incorporate information about alternative hypotheses into estimates of *p*(H|D) could be due to at least two distinct mechanisms. First, the reasoner may see the alternative sources of the data as irrelevant to the task of estimating *p*(H|D). Fiedler (2012) attributes such neglect to a meta-cognitive failure to consider the sampling or causal mechanisms that generate observed data. In a related view, Krynski and Tenenbaum (2007) suggest that people frequently fail to incorporate alternative sources of the data into their mental representation of judgment problems. We refer to these approaches collectively as "early stage neglect" accounts.

An alternative approach locates the failure to consider alternative hypotheses at a later stage in the judgment process. According to this view, even when people see the relevance of alternative hypotheses they fail to understand how the relevant statistical information should be combined to provide an estimate of *p*(H|D). In other words, people may fail to *utilize* statistical information relating to alternative hypotheses. Consistent with this utilization deficiency account, people often confuse the likelihood *p*(D|H) with the posterior probability *p*(H|D) (e.g., Gigerenzer & Hoffrage, 1995; Villejoubert & Mandel, 2002) and make errors in additively combining *p*(D|H) and *p*(D|¬H) (Tversky & Koehler, 1994).

Of course, these are not the only deviations from Bayesian norms in intuitive probability judgments. People also frequently neglect the base rate of the target hypothesis when estimating p(H|D) (e.g., Barbey & Sloman, 2007; Gigerenzer & Hoffrage, 1995). The current work however focuses primarily on why people fail to consider information about the alternative hypothesis.

The current studies aimed to test the early stage neglect and utilization deficiency accounts. Experiment 1 examined predictions that follow from the early stage neglect account. According to this account, people will generally fail to *spontaneously* look for information concerning alternative causes of the data when assessing the evidence for p(H|D), even when such a search requires minimal cognitive effort. However, because such neglect is thought to reflect a meta-cognitive bias rather than a capacity limitation (e.g., Fiedler, 2012) it should be possible to reduce neglect by providing explicit prompts about alternative sources of observed data.

Experiment 2 contrasted the predictions of the early stage neglect and utilization deficiency accounts. The neglect account suggests that when both the target and alternative hypotheses are stated explicitly and people are given the relevant likelihood statistics they will use this information appropriately. That is, they will update their beliefs about p(H|D) appropriately given the relative likelihoods of p(D|H) and p(D|¬H) (Fiedler, 2012; Juslin, Winman & Hansson, 2007). The utilization deficiency account predicts that judgment performance will show marked deviation from normative patterns, even when the hypothetical alternatives are well specified.

Experiment 1

Considerable evidence suggests that people are biased in favor of considering information that confirms a current focal hypothesis rather than seeking information about plausible alternatives (e.g., Beyth-Marom & Fischhoff, 1983; Doherty, et al., 1979). However, many of these studies confound search for information about the alternative hypothesis with the use or integration of statistical information. In many studies, for example, the likelihoods of the focal and alternative hypotheses are given as numerical probabilities. This means that participants have to integrate across these probabilities to arrive at an estimate of p(H|D). Moreover in many studies (e.g., Doherty, et al., 1979), participants are also expected to consider the base rates of the focal and alternative hypotheses.

To conduct a more direct test of the prediction that people will spontaneously fail to *search* for information about an alternative hypothesis, we devised a simplified task where searching successive Bayesian components involved minimal effort. The task was adapted from Doherty and Mynatt (1990). Participants took the role of a doctor seeking to diagnose whether a patient with a particular symptom (the datum, D) had a particular disease (the focal hypothesis, H). They were then presented with text descriptions of the four components of Bayes rule, p(H), p(¬H), p(D|H), p(D|¬H), and asked to choose the components that would be most helpful in arriving at a diagnosis.

Unlike many previous studies of neglect of the alternative hypothesis, no numerical values were presented and participants were not required to produce a numerical estimate of probability. Instead, the key outcome measure was the frequency with which information concerning the alternative hypothesis p(D|¬H) and/or p(¬H) was consulted during search. Following the early stage neglect hypothesis, it was predicted that participants would neglect information about the alternative in favor of information about the focal hypothesis, even when search for information about the alternative required minimal effort.

A second implication of the early-stage neglect approach is that the best way to improve statistical reasoning is to provide participants with more detailed information about the origins of the data that they are reasoning about (Fiedler, 2012). A number of studies have shown that providing explicit reminders about the existence of an alternative source of the data can increase the degree to which it is considered in diagnostic reasoning (e.g., Beyth-Marom & Fischhoff, 1983; Krynski & Tenenbaum. 2007). Given the strength of the bias towards the focal hypothesis, we expected that this bias could only be overcome when the alternative hypothesis was made just as salient as the focal hypothesis.

To test this prediction, we manipulated between groups the way in which the alternative hypothesis was presented in the search task. In the unspecified alternative condition, no alternative source of the observed symptom was mentioned. In this condition we expected that participants would generally neglect information about the alternative hypothesis in their search. In the generic alternative condition, it was suggested that the observed symptom could arise due to other diseases. In the specified alternative, a different disease was identified as an alternative cause of the observed symptom. Because this is the only condition where the salience of the alternative is likely to be similar to that of the focal hypothesis, we predicted that this would be the only condition where we observe increased search for information relating to the alternative hypothesis.

Method

Participants. A total of 903 participants (39% female; M_{AGE} = 30.39 years) from the USA were recruited via Amazon Mechanical Turk and were paid 50c US. Three were excluded because they failed to respond correctly to the attention screening item (described below). Participants were randomly allocated to one of three conditions.

Design and Procedure. Participants in all conditions were told to imagine that they were a doctor whose task was to identify information that would assist in diagnosing whether a patient who showed a particular symptom ("red rash on their fingers") had a target disease ("Buragamo"). Participants were instructed to select only those options that were "necessary for the diagnosis". 10 s later four search options appeared. These were the base rate of the target hypothesis $p(H)$, the base rate of the alternative hypothesis $p(\neg H)$, the likelihood of the data (i.e., red rash) given the target hypothesis $p(D|H)$, and the likelihood of data given an alternative hypothesis $p(D|\neg H)$. The four options were presented in rectangles of different colors, arranged in two rows of two. The location of each option was randomized for each participant. Participants clicked on a rectangle to indicate their choice. Once selected the rectangle changed colour. A previously selected search option could be de-selected by clicking on it a second time. Participants were required to select at least one option and were free to select between one and four options. Once selections were finalized participants advanced to an attention screening question, in which four unlabeled colored rectangles were presented. Participants were required to click on the two rectangles with the same color.

Three experimental conditions differed in their specification of an alternative hypothesis (i.e., a different diagnosis). In the unspecified alternative condition only the target disease was mentioned in the instructions and search options and the alternative hypothesis was "does not have Buragamo". The generic alternative condition was similar except that it was noted in instructions that the symptom of red rash "could be caused by a number of diseases". In the specified alternative condition participants were told that the disease "Terragaxis" could also cause the red rash and "has Terragaxis" was specified as the alternative hypothesis in the search options.

Results

The four search options yielded a total of 15 possible search patterns (ignoring search order). The proportion of participants in each condition whose searches conformed to each pattern is given in Table 1. The Table shows that all groups saw $p(D|H)$ as important and were likely to search for this information either as their only search choice or in combination with other options. In the unspecified alternative and generic alternative groups the three most common strategies in descending order of frequency were $p(D|H)$, $p(D|H) + p(D|\neg H)$, and $p(H) + p(D|H)$. Those in the specific alternative condition showed a different pattern. In this case the modal choice was $p(D|H) + p(D|\neg H)$, selected by more than half the participants, with $p(D|H)$ the next most common choice. The percentage of participants selecting $p(D|H) + p(D|\neg H)$ in the specified alternative condition (50.7%) was reliably higher than in the other two conditions; unspecified alternative (22.7%, $z = 7.08$, $p < .001$), generic alternative (25.7%, $z = 6.29$, $p < .001$).

A more inclusive test of differences in consideration of the alternative hypothesis was carried out by aggregating search patterns according to whether or not they included $p(D|\neg H)$ and/or $p(\neg H)$. The proportion of searches that included the alternative hypothesis was significantly higher in the specified alternative group (70.5%) than in the other two conditions (unspecified: 41.1%; generic: 41.9%), $\chi^2(2, N=900) = 68.85$, $p < .001$.

Although the main focus was on search for information about the alternative hypothesis, we also examined group differences in consideration of the base rate of the focal hypothesis. The proportion of searches that included $p(H)$ was significantly lower in the specified alternative condition (18.5%) than in the other two conditions ($M = 32.1\%$), $\chi^2(2, N=900) = 22.61$, $p < .001$.

Table 1. Proportion of participants in each condition following each search pattern (with modal choice in **bold**)

Search Patterns	Unspecified alternative n = 299	Generic alternative n = 303	Specified alternative n = 298
$p(H)$	0.020	0.013	0.023
$p(\neg H)$	0.003	0.000	0.007
$p(D\|H)$	**0.344**	**0.370**	0.194
$p(D\|\neg H)$	0.050	0.066	0.027
$p(H), p(\neg H)$	0.000	0.003	0.054
$p(H), p(D\|H)$	0.224	0.198	0.077
$p(H), p(D\|\neg H)$	0.040	0.036	0.013
$p(D\|H), p(D\|\neg H)$	0.227	0.257	**0.507**
$p(\neg H), p(D\|H)$	0.003	0.013	0.013
$p(\neg H), p(D\|\neg H)$	0.003	0.007	0.003
$p(H), p(\neg H), p(D\|H)$	0.000	0.000	0.003
$p(H), p(\neg H), p(D\|\neg H)$	0.000	0.000	0.003
$p(H), p(D\|H), p(D\|\neg H)$	0.074	0.033	0.010
$p(\neg H), p(D\|H), p(D\|\neg H)$	0.003	0.000	0.007
$p(H), p(\neg H), p(D\|H), p(D\|\neg H)$	0.007	0.003	0.057

Discussion

In this study participants searched for information that they believed to be relevant for diagnosing a particular disease (the focal hypothesis) given observed data. The key question was whether people would spontaneously search for information about an alternative cause of the data (the alternative hypothesis). The first major finding was that despite the minimal effort required for information search, most participants neglected the alternative hypothesis. Less than half of those in the unspecified and generic conditions searched for any information about the alternative hypothesis.

Such neglect is consistent with a range of previous findings showing a preference for examination of statistical information about the focal hypothesis over the alternative (e.g., Beyth-Marom & Fischhoff, 1983; Doherty, et al., 1979). Notably the current results show that such neglect persists when demands on attention and memory are minimized and no numerical calculation was required.

This study also clarified the conditions under which this bias can be overcome. Providing a specific but not a generic alternative hypothesis led a majority of participants to search for *p*(D|¬H) and/or *p*(¬H).

As in many previous studies (e.g., Gigerenzer & Hoffrage, 1995) only a minority of people believed base rate information was relevant to evaluating *p*(H|D). An unexpected finding was that increased search for information about the alternative hypothesis in the specified alternative condition appeared to come at the expense of information about the base rate of the focal hypothesis. We return to this issue in the General Discussion.

Experiment 2

Experiment 1 found that in line with the early stage neglect account, people do not spontaneously consider the alternative hypothesis when evaluating *p*(H|D). Experiment 2 examined whether people understand the implications of *p*(D|¬H) when the alternative hypothesis is clearly specified. The utilization deficiency account predicts that people will fail to understand the implications of statistical information associated with the alternative hypothesis even when this hypothesis is clearly specified. Alternately, the early-stage neglect account suggests that once the relevance of the alternative hypothesis is established, people will factor the relevant likelihood statistics into their estimation of *p*(H|D).

To test these accounts participants were given a diagnostic problem in which they were sequentially presented with numerical information corresponding to the major components of Bayes' theorem. In stage 1 they were given the base rate of a focal hypothesis (the disease Buragamo) and asked to estimate *p*(H|D). In stage 2 they were presented a new datum (red rash) and with the likelihoods of *p*(D|H) and *p*(D|¬H). Participants were asked to re-estimate *p*(H|D) in the light of this information.

The key question was whether participants used the likelihood information in a normative way when revising probability estimates. Normatively, when the focal and alternative hypotheses are equally likely then no revision in initial estimates should take place. If the likelihood of *p*(D|H) is substantially greater than *p*(D|¬H) then estimates should be revised upwards. If the likelihood of *p*(D|H) is substantially lower than *p*(D|¬H) then estimates should be revised downwards. Table 2 shows the details of the Bayesian predictions (cf. Beyth-Marom & Fischhoff, 1983).

According to early-stage neglect accounts, the provision of the statistical details of *p*(D|¬H) should in itself increase the salience of the alternative hypothesis at an early stage of problem representation. Hence, the way that the alternative hypothesis is framed (e.g., unspecified vs. specified) should have less impact on evaluating *p*(H|D) than was found in Experiment 1. To test this prediction we again manipulated the framing of the alternative hypothesis.

Method

Participants. A total of 959 US participants (45% female; M_{AGE} = 32.98 years) were recruited via Amazon Mechanical Turk and were paid $1.00 US. Eleven were excluded because they failed the attention check, gave zero probability estimates at either stage 1 or 2, or reported having previously completed a similar study. Participants were randomly allocated to eight experimental conditions of roughly equal size (*n* = 117-123).

Table 2. Summary of the likelihood conditions in Experiment 2 and Bayesian predictions.

Likelihood Condition	Stage 1 *p*(H)	Stage 2 *p*(D|H)	Stage 2 *p*(D|¬H)	*p*(H|D) from Bayes' rule
High/High	0.8	0.9	0.9	0.8
Low/Low	0.8	0.1	0.1	0.8
High/Low	0.8	0.9	0.1	0.97
Low/High	0.8	0.1	0.9	0.31

Design and Procedure. The experiment was conducted in two stages. In stage one, participants were told that they were to take the role of a doctor and to use the information given to assess the probability that a new patient X has the disease Buragamo. They were then given the base rate of the disease ("80% of patients that you have seen had Buragamo. The remaining patients did not have Buragamo") and asked "What do you think is the probability that Patient X will have Buragamo?" Answers were given as a percentage.

In stage two participants were told that the target patient had a red rash and that the following was known from medical records: "X% of patients that you have seen WITH Buragamo have a red rash and that Y% of patients that you have seen WITHOUT Buragamo [WITH another disease Terragaxis] have a red rash. Four different combinations of the likelihoods of *p*(D|H) and *p*(D|¬H) were administered to different groups (high/high, low/low, high/low, low/high). Table 2 shows the likelihood figures presented in each condition.

These conditions were crossed with a manipulation of the framing of the alternative hypothesis to give a total of eight between-subjects conditions. In the unspecified condition, the alternative was described as the absence of the focal disease. In the specified condition the alternative was the disease Terragaxis (see alternative instructions given in square brackets above). Participants were then given their stage 1 probability estimate and asked "What do you now think is the probability that Patient X has Buragamo?" Answers were again given as a percentage.

Results and Discussion

At stage 1, 78% of subjects gave the normatively correct estimate of "80" based solely on the base rate of the focal hypothesis. However there was a marginal trend for probability estimates in the specified alternative condition to be higher (M = 75.89) than those in the unspecified condition (M = 74.25), $F(7, 940) = 2.00$, $p = .05$. Hence in analyses of stage 2 estimates, individuals' stage 1 estimates were used as a covariate.

The key dependent measure was the change in estimates of p(H|D) between stages 1 and 2. Because the predicted direction of change differed for high/high vs. high/low likelihood and low/low vs. low/high likelihood conditions, these conditions were analyzed separately. If people considered the likelihoods of both the target hypothesis p(D|H) and the alternative p(D|¬H), there should have been little change in estimates from stage 1 to stage 2 in the high/high condition but an increase in estimates in the high/low condition (see Table 2). A 2 (likelihood condition) x 2 (framing of alternative) analysis of covariance (ANCOVA) revealed a reliable difference in the change in estimates in the high/low as compared with high/high conditions, $F(1,473) = 91.96$, $p < .001$. Figure 1 shows that estimates in the high/high condition did not change from stage 1 to 2. However estimates in the high/low condition were revised upwards, consistent with Bayesian predictions. The specification of the alternative hypothesis did not affect change in the probability estimates (F's < 1.0).

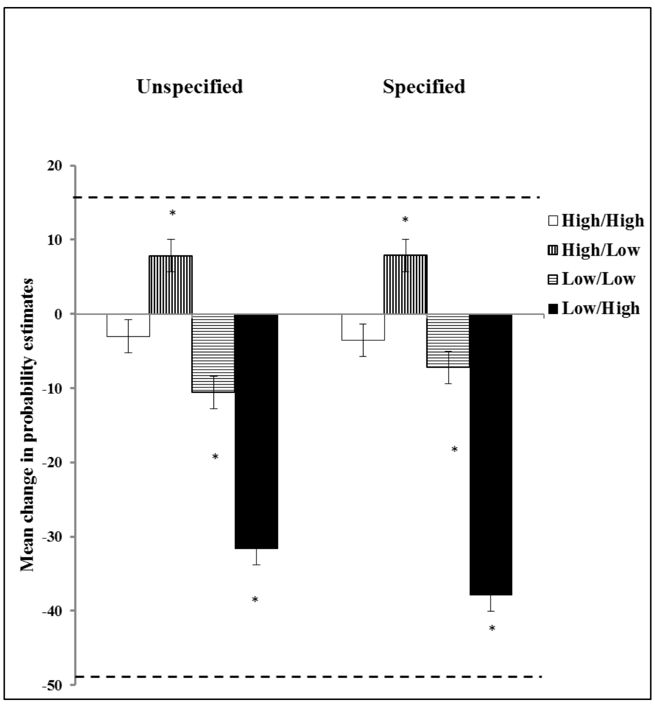

Figure 1. Mean change in probability estimates from stage 1 to stage 2 in each experimental condition.
Note: Dashed lines show normative levels of belief revision for the high/low (upper) and low/high (lower) conditions. * denotes significantly different from zero change, $p < .001$

If people were considering information about both the target and alternative hypotheses, then there should have been little change in the low/low condition from stage 1 to 2 but a decrease in estimates in the low/high condition. Figure 1 shows that there was an unexpected decrease in probability estimates in the low/low condition. Nevertheless, consistent with Bayesian predictions, the magnitude of change in estimates from stage 1 to stage 2 was larger in the low/high than the low/low condition, $F(1,465) = 139.07$, $p < .001$. There was a further interaction between this effect and the framing of the alternative, $F(1,465) = 4.88$, $p = .03$. As shown in the Figure, the difference in change between the low/low and low/high conditions was larger when an alternative cause was specified.

Bayes' rule predicts that absolute change in estimates in the high/low condition should be smaller (normatively ≈ 17.29%) than in the low/high condition (normatively ≈ 49.23%). This prediction was also confirmed, $F(1,473) = 231.51$, $p < .001$. However Figure 1 shows that updating was conservative with change in estimates generally smaller than the values given by Bayes' theorem.

These results show that people generally understood the implications of the relative likelihoods of the focal and alternative hypotheses. Their use of this likelihood information conformed to Bayesian principles, at least at a qualitative level. The results stand in stark contrast to Experiment 1 where the majority of participants ignored information about the alternative hypothesis.

General Discussion

These experiments aimed to clarify why people often fail to incorporate information about the alternative hypothesis in judgments under uncertainty. Two competing accounts were proposed. The early stage neglect hypothesis suggests that the core problem is that people often do not see the relevance of the alternative hypothesis when developing a representation of the problem at hand (Fiedler, 2012; Krynski & Tenenbaum, 2007). An alternative view is that people fail to understand the implications of statistics associated with the alternative hypothesis even when the alternatives are stated explicitly (Tversky & Koehler, 1994; Villejoubert & Mandel, 2002).

The results of the two experiments strongly support the first account. Consistent with the notion of early stage neglect, people frequently failed to consult information about the alternative hypothesis when asked to evaluate p(H|D) (Experiment 1). This was the case even though the tasks were designed to minimize the cognitive effort involved in considering the alternative hypothesis. In contrast, when the likelihoods of the data given the focal and alternative hypotheses were stated explicitly (Experiment 2), most participants used this information in a manner that was broadly consistent with a normative approach. In sum, people generally showed little spontaneous recognition of the relevance of the alternative hypothesis to the evaluation of p(H|D) but were able to use statistical information about each hypothesis appropriately when this information was supplied.

A further important finding in Experiment 1 was that consideration of the alternative hypothesis could be increased by labeling it in a way that increased its salience relative to the focal hypothesis (cf. Beyth-Marom & Fischhoff, 1983). This suggests that neglect of the alternative represents a misunderstanding of the structure of the problem rather than a capacity limitation in simultaneously considering the alternatives.

Our findings are relevant to accounts of judgment and reasoning which propose that people focus on one hypothesis at a time and avoid the consideration of uncertain alternatives. This is exemplified in Evans' (2006) *singularity principle* which states that "people construct only one mental model at a time with which to represent a hypothetical situation" (p. 379).

The current work suggests that this principle is only partly true. Consistent with the singularity principle we found that neglect of the alternative largely occurs in the initial representation of the problem. However, our work goes beyond singularity by showing that such neglect can be reduced by making the alternative more salient so that is seen as a competing causal explanation of the observed data. Moreover, we have shown that people are readily able to *use* the likelihood statistics associated with multiple hypotheses when estimating $p(H|D)$.

A surprising finding in Experiment 1 was that an increase in search for information about the alternative hypothesis was accompanied by a reduction in search for the focal base rate. Identifying the mechanisms behind this tradeoff is an important issue for future research. A possible explanation is suggested by previous work examining training interventions targeting base rate neglect. Fischhoff and Bar-Hillel (1984) found that such interventions often lead to increased attention to the base rate in both problems where base rate use is normative as well as problems where the normative strategy is to ignore the base rate. In other words, training increased attention to certain aspects of judgment problems but did not necessarily improve understanding of the contexts in which it was appropriate to use or ignore these components. Likewise in Experiment 1 priming attention to the alternative hypothesis may have led to neglect of other relevant components of the problem.

These studies show that the main locus of neglect of the alternative hypothesis in judgments under uncertainty is at the early stage of representing the problem. We found evidence of neglect of information about the alternative in tasks where minimal effort was required to examine such information and numerical estimates of probability were not required. In contrast, when statistical information about an alternative hypothesis was provided, the qualitative pattern of people's probability judgments conformed to Bayesian prescriptions. These findings are consistent with accounts of judgment under uncertainty which emphasize that people are often insensitive to alternative ways that observed data might be generated but are relatively accurate when utilizing statistical information given to them (e.g., Fiedler, 2012).

Acknowledgments

This work was supported by Australian Research Council Discovery Grant DP120100266 to the first and third authors and a Future Fellowship FT110100151 to the third author. We thank Kate Blundell and Kelly Jones for help with programming and running the experiments.

References

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*, 241-254.

Beyth-Marom, R., & Fischhoff, B. (1983). Diagnosticity and pseudodiagnosticity. *Journal of Personality and Social Psychology*, *45*, 1185-1195.

Doherty, M. E., & Mynatt, C. R. (1990). Inattention to $p(H)$ and to $p(D|\neg H)$: A converging operation. *Acta Psychologica*, *75*, 1-11.

Doherty, M. E., Mynatt, C. R., Tweney, R. D., & Schiavo, M. D. (1979). Pseudodiagnosticity. *Acta Psychologica*, *43*, 111-121.

Dougherty, M., Thomas, R., & Lange, N. (2010). Toward an integrative theory of hypothesis generation, probability judgment, and hypothesis testing. *The Psychology of Learning and Motivation*, *52*, 299-342.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press

Evans, J. S. B. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, *13*, 378-395.

Fiedler, K. (2012). Meta-cognitive myopia and the dilemmas of inductive-statistical inference. *The Psychology of Learning and Motivation*, *57*, 1–46.

Fischhoff, B., & Bar-Hillel, M. (1984). Focusing techniques: A shortcut to improving probability judgments? *Organizational Behavior and Human Performance, 34*, 175-194.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684-704.

Hayes, B. K., Hawkins, G. E., Newell, B. R., Pasqualino, M., & Rehder, B. (2014). The role of causal models in multiple judgments under uncertainty. *Cognition*, *133*, 611-620.

Juslin, P., Winman, A., & Hansson, P. (2007). The naïve intuitive statistician: A naïve sampling model of intuitive confidence intervals. *Psychological Review*, *114*, 678-703.

Krynski, T. R., & Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology: General*, *136*, 430–450.

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*, 547-567.

Villejoubert, G., & Mandel, D. R. (2002). The inverse fallacy: An account of deviations from Bayes' theorem and the additivity principle. *Memory & Cognition*, *30*, 171-178.