THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2012, 65 (3), 541–552

Psychology Press
Taylor & Francis Group

Relevance-driven information search in "pseudodiagnostic" reasoning

Gaëlle Villejoubert and Frédéric Vallée-Tourangeau

Department of Psychology, Kingston University, Kingston upon Thames, Surrey, UK

When faced with two competing hypotheses, people sometimes prefer to look at multiple sources of information in support of one hypothesis rather than to establish the diagnostic value of a single piece of information for the two hypotheses. This is termed pseudodiagnostic reasoning and has often been understood to reflect, among other things, poor information search strategies. Past research suggests that diagnostic reasoning may be more easily fostered when participants seek data to help in the selection of one of two competing courses of action as opposed to situations where they seek data to help infer which of two competing hypotheses is true. In the experiment reported here, we provide the first empirical evidence demonstrating that manipulating the relevance of the feature for which participants initially receive information determines whether they will make a nominally diagnostic or pseudodiagnostic selection. The discussion of these findings focuses on implications for the ability to engage in diagnostic hypothesis testing.

Keywords: Hypothesis testing; Bayesian reasoning; Pseudodiagnosticity.

A sales manager advertises a new position for a sales assistant. After reviewing the curriculum vitae of the candidates, she selects two promising applicants, Ms. A. and Ms. B. The manager initially leans towards the first candidate, Ms. A. and discovers that she successfully completed 70% of her sales transactions in the last month in her previous position. The manager must now engage in inductive reasoning: She needs to collect more information in order to decide whether Ms. A or Ms. B. is the best candidate for the job. A long tradition of psychological research suggests that her search for information will be driven by a need for evidence confirming the hypothesis she is entertaining. Thus, if, at this point, the manager believes Ms. A. *is* the best candidate, she may naturally seek more information about Ms. A rather than checking Ms. B.'s sale performance. Yet, this strategy is potentially shortsighted: Ms. B. could well have outperformed Ms. A. in terms of sales, in which case, seeking more information about Ms. A. could be misguided and lead to the employment of a lesser able candidate.

More generally, the diagnosticity of a datum D for purposes of discriminating between a given

Correspondence should be addressed to Gaëlle Villejoubert or Frédéric Vallée-Tourangeau, Department of Psychology, Kingston University, Kingston upon Thames, Surrey, KT1 2EE, UK. E-mail: g.villejoubert@kingston.ac.uk or f.vallee-tourangeau@kingston.ac.uk
We would like to thank C. R. McKenzie, one anonymous reviewer, and J. D. Nelson for helpful comments and guidance on previous versions of this manuscript. We would also like to thank R.V. Gow for the initial suggestion of formulating a scenario involving high schools.

http://www.psypress.com/qjep
http://dx.doi.org/10.1080/17470218.2011.617830

hypothesis X (H_X) and an alternative hypothesis Y (H_Y) is defined in terms of the ratio of the probability that D is observed given that H_X is true, $P(D|H_X)$ and the probability that D is observed given that an alternative hypothesis, Y, is true, $P(D|H_Y)$. Hence, diagnosticity can only be assessed from the perspective of multiple hypotheses. Psychologists have traditionally assumed that the likelihood ratio in Bayes's theorem was a metric of the diagnosticity of information (e.g., Crupi, Tentori, & Lombardi, 2009; Doherty, Mynatt, Tweney, & Schiavo, 1979; Tweney, Doherty, & Kleiter, 2010).

Information search and diagnosticity

Early research examining how people gather information in order to make inferences suggested that they did not fully appreciate that diagnosticity is defined in terms relative to at least two hypotheses, not just one (Beyth-Marom & Fischhoff, 1983; Doherty et al., 1979; Kern & Doherty, 1982 ;cf. Trope & Bassok, 1982, 1983). Mynatt, Doherty, and Dragan (1993) argued that people's hypothesis testing process is predominantly concerned with gathering evidence in favour of one hypothesis, rather than determining the diagnosticity of any given piece of information for multiple hypotheses. In one of their reasoning scenarios, participants were asked to determine which of two cars, X or Y, their "sister" purchased. They were told about *two features characterizing this car*, its petrol consumption (D_1: "25 miles per gallon, mpg") and its mechanical reliability (D_2: "no major mechanical problems in the first two years of ownership"). In addition, participants were given *an anchoring piece of information*: "65% of car Xs do over 25 mpg", or $P(D_1|H_X) = .65$. Participants were then asked to choose which of the following three pieces of statistical information would help them decide which type of car was owned by their sister (the participants did not see the information presented here in brackets; the labels "row", "column", and "diagonal" values refer to the values arranged in a 2 × 2 array with H_X and H_Y as column headers and D_1 and D_2 as row headers; see Table 1).

1. The percentage of car Ys that do over 25 mpg [$P(D_1|H_Y)$, row value].
2. The percentage of car Xs that have had no major mechanical problems for the first two years of ownership [$P(D_2|H_X)$, column value].
3. The percentage of car Ys that have had no major mechanical problems for the first two years of ownership [$P(D_2|H_Y)$, diagonal value].

In Experiment 1 of Mynatt et al. (1993), 60% of the participants chose to learn about the percentage of car Xs with good mechanical reliability in order to determine the identity of the car (the column value), while only 26% chose to know the percentage of car Ys that did 25 mpg (the row value).

Up until recently, scholars had assumed that selecting the row value was the only optimal strategy in this task as it allows computing a likelihood ratio, which informs the diagnosticity of the feature associated with the anchoring value. This choice was therefore customarily referred to as the "diagnostic" choice. By contrast, choices to uncover the column or the diagonal values were considered to reflect poor information search strategies, since neither value would allow one to evaluate the diagnosticity of the anchoring feature. These choices were consequently labelled "pseudodiagnostic" (e.g., Mynatt et al., 1993). We note that this analysis of the task has been recently

Table 1. *Information set in "pseudodiagnostic" tasks*

	Hypotheses	
Feature	H_X	H_Y
D_1	Visible anchoring value, $P(D_1|H_X)$	Hidden row value, $P(D_1|H_Y)$
D_2	Hidden column value, $P(D_2|H_X)$	Hidden diagonal value, $P(D_2|H_Y)$

challenged. Crupi et al. (2009) argued that, under some assumptions that could be adopted by participants, nominally pseudodiagnostic information is more useful than nominally diagnostic information. Others, however, have since contended that this challenge only holds when some—debatable—conditions apply, whereas the initial analysis holds under less restrictive assumptions (Tweney et al., 2010). We return to this recent controversy in the Discussion. Meanwhile, we adopt a more neutral labelling of the choices as "row", "column", and "diagonal" choices.

Feature relevance and initial values in action problems

The goal of the inferential task is another variable that influences the degree to which people make diagnostic search choices. Mynatt et al. (1993) distinguished between *inference* and *action* problems. The car example discussed above, they argued, represents an inference problem. The car has already been purchased and is owned by someone; the goal is to determine whether it is a car X or a car Y. In effect the problem is a categorization inference, and in principle the categorization can be true or false. In contrast, an action problem is one where courses of actions are evaluated and correspond to the two hypotheses of a traditionally formulated inference problem. One course of action might be better than the other, but the decision cannot in principle be evaluated in terms of whether one action is true and the other false. In a separate experimental condition, Mynatt et al. (1993) asked participants to imagine *buying* a car (emphasis added), considering car X or car Y, and told them they were "concerned about . . . petrol consumption and mechanical reliability" (p. 768). Participants were then given the same anchoring piece of information ("65% of car Xs do over 25 mpg") and were asked to choose one among the same three pieces of information in order to help them decide which car to buy (see Options 1 through 3 above).

In that action problem, 52% chose the piece of information that could determine the petrol consumption of car Ys (the row choice), and 41% chose the piece of information that could determine the mechanical reliability of car Xs (a column choice). To explain the high proportion of row choices in action problems, Mynatt et al. (1993) propose that the choice among the three alternatives is determined by the datum that bears more utility for each individual participant: "Precisely how many subjects will select [the row choice] will depend on the content of a given problem and subjects' *idiosyncratic* utility function and decision strategies" (pp. 765–766, emphasis added). An important distinction should be made here between the utility of the *numerical anchor*, P(D|H), and the utility of knowing that the *feature*, D, has been observed in relation to the decision to be made. According to Savage's (1954) expected utility theory, the utility of information is contingent on its potential to foster better decisions. According to the classical view on pseudodiagnosticity (e.g., Mynatt et al., 1993; Tweney et al., 2010), the numerical *value* uncovered with a column choice bears zero utility for the decision since it cannot help improving it. This does not preclude, however, that the two *features* about which participants receive statistical information—D_1 and D_2—may be deemed more or less useful for the decision at hand. To avoid confusion, we refer to this second type of utility as the feature *relevance* for the decision—that is, the extent to which it is has some bearing or importance for the decision at hand.

On this account, those who consider petrol consumption to be more relevant than mechanical reliability for deciding which course of action should be pursued would be motivated to establish the petrol consumption of car Ys and hence chose to inspect the row value. In contrast, those who are more concerned about mechanical reliability should seek information about car Xs' mechanical reliability, the column choice.

The authors, however, did not manipulate explicitly the perceived relevance of the two features characterizing each alternative (e.g., petrol consumption and mechanical reliability in the car scenario), nor did they seek to assess how relevant their participants believed these features to be. Moreover, there is conflicting evidence showing that action problems may not necessarily promote row choices. Maggi, Butera, Legrenzi, and Mugny

(1998, Experiment 1) asked participants to imagine having to choose between two cars or two political candidates. These authors found over 60% choices to be column choices. There was, however, an important methodological difference between their task and that of Mynatt et al. (1993): Maggi et al. (1998) presented participants with *four* possible pieces of information to choose from for each alternative (e.g., the car price, reliability, fuel consumption, and performance). In addition, the authors found that people tended to prefer row selections when the anchoring feature concerned a characteristic they believed to be important (e.g., the price of a car or the competence of a political candidate). In light of these incongruous findings, the main goal of the present experiment was thus to establish whether the relevance of the anchoring feature for the decision at hand would cause individuals to make a row or column selection.

Another important difference between inference and action problems outlined by Mynatt et al. (1993) is the role of the initial $P(D_1|H_X)$ value of the anchoring feature. According to the authors, in inference problems this initial value could be a cue to the truth value of H_X and, as such, dictate participants' information search. By contrast, the authors predicted and found that this initial value would not affect choices in action problems since, in those situations, information search would solely be determined by the relevance of the anchoring feature D_1. The authors, however, tested this prediction by comparing relatively narrow values, hovering modestly below and above the 50% mark (viz. 35% vs. 65%). It is therefore reasonable to assume that participants who believed, for example, that petrol consumption was *the* most important attribute for a new car would always wonder if car Ys outperformed car Xs, even upon learning that 65% of car Xs did over 25 mpg. However, this does not necessarily imply that participants' information search strategy will never be affected by the $P(D_1|H_X)$ value in action problems. It is not implausible to expect, for example, that participants may no longer seek to uncover the row value when told that 95% of car Xs do over 25 mpg. In this case, the $P(D_1|H_X)$ value could be deemed a sufficient cue to the superiority of H_X. Consequently, under such circumstances, participants might then be more interested in learning more about car Xs than in finding out whether car Ys outperform car Xs. A secondary goal of the present experiment was therefore to test whether an increase in column choices would be observed when the $P(D_1|H_X)$ value is deemed satisfactory in action problems.

The present experiment

Mynatt et al.'s (1993) account suggests that information search in action problems will be driven by relevance considerations. Accordingly, when the first or anchoring feature given is deemed less relevant than the second feature, column choices should be observed. This is because participants would be more interested in determining the value of the more important feature for the focal hypothesis H_X, in this instance the nonfocal feature, D_2. If, however, the anchoring feature is deemed highly relevant to determine whether the focal hypothesis should be favoured, a higher proportion of row choices should be observed, reflecting the participants' motivation to receive statistical information about this feature for the alternative hypothesis. Moreover, if information search is driven by relevance considerations, the $P(D_1|H_X)$ value should *only* influence the rate of row choices when D_1 refers to a high-relevance feature.

Method

Participants

Participants were recruited by third-year psychology students at the University of Toulouse, France, as a course requirement. Each student made a list of several men and women who were older than 18 years and not studying psychology, randomly drew one man and one woman from his or her list, and asked them to take part in a general survey, which included the present study. Of the 1,040 participants in the final sample (520 men, 520 women; mean age = 31.37 years, $SD =$ 13.24), 11% had completed graduate school, 53% had an undergraduate education, 20% had graduated from high school only, and the remaining

16% had not graduated from high school. The sample included a large proportion of students (40%), but also working professionals (51%) and retired or unemployed individuals (8%). The survey was conducted in French.

Design and procedure

The "pseudodiagnostic" task was part of a six-page survey and was embedded among a number of different tasks and questionnaires. The present experiment used a 2 × 2 between-subjects design. The independent variables were the relevance of the anchoring feature given (high or low) and its numerical value (high or low). Participants were randomly allocated to one of the resulting four conditions. The reasoning task was presented on page 3 of the survey and was described as follows (see the Appendix): Participants were asked to imagine a high-school student seeking to know more about two possible high schools to attend, high school X and high school Y. This student was said to look for a high school, which offers social and cultural events as well as a solid education since he would like to gain access to a top university. Participants were then told the student enquired about both high schools and found one piece of statistical information about high school X. Half the participants were then given a piece of statistical information about the number of parties organized by high school X while the remaining participants were told about the percentage of students graduating from high school X who gained access to a top university. Within each condition, the piece of statistical information was characterized by either a low value (one party organized every month; 20% accessed a top university) or a high value (one party organized every week; 80% accessed a top university). Following the presentation of this first piece of information, participants were asked to select among three additional pieces of statistical information the student should consult before making a choice, assuming the student could only have access to only one additional piece of information: (a) They could advise the student to consult the value of the same piece of information concerning high school Y (the row choice); (b) they could advise consulting the value of the alternative piece of information for high school X (i.e., consulting the frequency of parties if told about access rates or consulting access rates if told about parties—the column choice); and (c) they could advise consulting the value of the alternative piece of information for high school Y (a diagonal choice). The order of the three choice options was constant across all four experimental conditions.

On page 5 of the survey, after completing an intermediary task, participants were asked to consider the high-school student task again and rate the importance of (a) parties organized by a high school that aims to prepare students to gain entrance to a top university and (b) gaining entrance to a top university upon leaving high school. Both ratings were recorded on an 8-point scale ranging from 1 (absolutely not important) to 8 (extremely important).

Results

Importance ratings

Before analysing row choices, we conducted manipulation checks on the importance ratings to verify that participants did believe that gaining entrance to a top university (TU) was more important than the number of parties (P) they could attend. Participants gave higher ratings to gaining entrance to a top university than to the number of parties organized by the high school in the low ($M_{TU} = 6.88$ vs. $M_P = 3.95$) and high anchor value condition ($M_{TU} = 6.84$ vs. $M_P = 3.71$) when D_1 concerned parties. The same pattern in the low ($M_{TU} = 6.74$ vs. $M_P = 3.85$) and high anchor value condition ($M_{TU} = 6.59$ vs. $M_P = 3.99$) was observed when D_1 concerned access to a top university. In a 2 (feature relevance) × 2 (low or high anchor value) × 2 (importance of parties vs. entry to a top university) mixed analysis of variance (ANOVA), the main effect of the feature relevance was not significant, $F < 1$, nor was the main effect of the anchor value, $F(1, 1036) = 1.1$. However, the importance ratings of being admitted to a top university were significantly higher than the ratings of the number of parties held during the year, $F(1, 1036) = 1{,}406$, $MSE = 3.08$, $p < .001$, $\eta_p^2 = .58$. Finally, none of the interactions were significant, largest nonreliable $F(1, 1036) = 3.32$. These results thus establish that participants considered that being admitted to a top university was a more important feature of a high school than the number of parties. Next, we examine whether information search strategy will vary as a function of the importance of the feature associated with the anchoring value.

Choice preferences

The frequency of row, column, and diagonal choices in the four experimental conditions are plotted as percentages in Figure 1. A number of patterns characterized these data. First, the number of row choices was very high in the scenarios with the high-relevance anchoring feature (D_1 = access to top university) but very low in the scenarios with the low-relevance anchoring feature (D_1 = number of parties). Second, within the scenarios using the low-relevance feature, variations in the anchor value of $P(D|H_X)$ appeared to have had no impact on the choice frequencies of the row and column options. In turn, the percentage of column choices increased from 9% to 20% when D_1 concerned access to a top university, and $P(D|H_X)$ equalled .8. Finally, the frequency of diagonal choices was higher in the scenarios when the number of parties anchored the focal hypothesis than when access to a top university did (in the latter the percentage of participants who switched did not exceed 2%).

The choice frequencies were modelled using a three-way model selection log–linear analysis. The fully saturated model (including all one-way and two-way effects, as well as the three-way interaction) was retained since deleting the three-way interaction term from the model led to a significantly lower goodness-of-fit index, $\Delta df = 2$, $\Delta G^2 = 12.21$, $\Delta p = .002$. The significant parameters that describe the final model, as determined by the log–linear analysis, are presented in Table 2.

Separate chi-square tests were conducted to further explore the three-way interaction. First the frequency of row choices in the parties versus entrance to top university scenarios were compared: There were significantly more row choices with the scenarios with the high-relevance feature than with the scenarios with the low-relevance feature, $\chi^2(3, N = 961) = 465$, $p < .001$. Second, the frequency of the different choices within the scenarios using number of parties as the anchoring feature was analysed: These frequencies did not significantly differ as a function of the number of parties, $\chi^2(2, N = 516) = 1.56$, $p > .05$. Third, a similar analysis

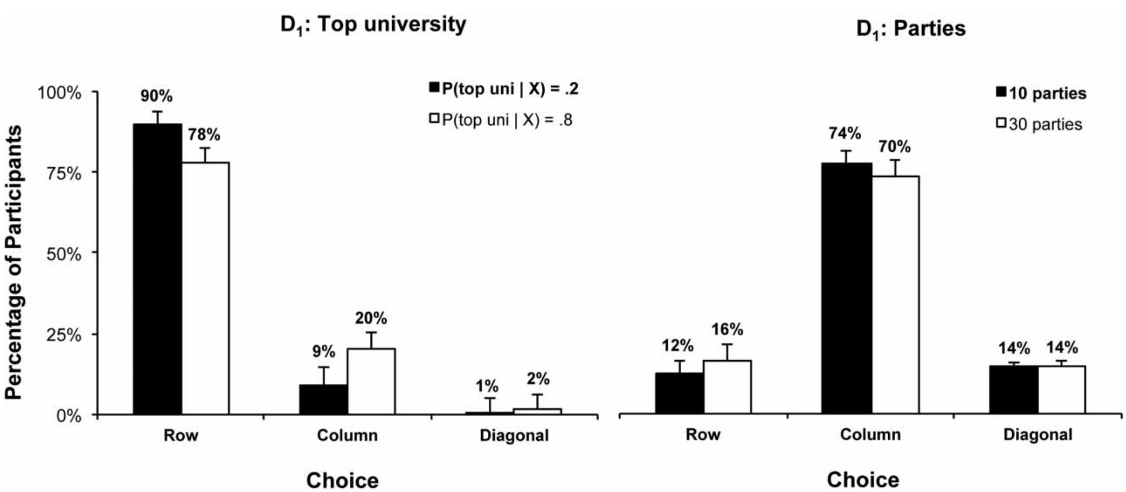

Figure 1. *Percentage of participants making a row, column, or diagonal choice for the scenarios using a high-relevance anchoring feature (proportion gaining entrance to a top university) and for the scenarios using a low-relevance anchoring feature (number of parties a year). Error bars represent 95% confidence intervals.*

Table 2. *Log–linear parameter estimates and values for choice, relevance, and anchor*

Effect		*Coefficient*	Z
Choice × Relevance × Anchor	Choice = row	−0.200	−2.434*
Choice × Relevance	Choice = row	1.261	15.369***
	Choice = column	−0.469	−5.677***
Choice	Choice = row	0.702	8.557***
	Choice = column	0.627	7.589***
Relevance		−0.352	−4.801***

Note: Saturated model. Choice: 1 = row, 2 = column, 3 = diagonal. Relevance: 1 = high, 2 = low. Anchor: 1 = high, 2 = low.
*$p < .05$. ***$p < .001$.

was conducted within scenarios using access to a top university as the anchoring feature: Choice frequencies varied significantly as a function of the anchor value of $P(D|H_X)$, $\chi^2(1, N = 524) = 14.0$, $p < .002$. Participants were more likely to choose to know the number of parties held at high school X if 80% of its students were admitted to a top university.

Individual differences

If information search is driven by idiosyncratic judgements of the relevance of the anchoring feature, then those who made row choices should also believe that D_1 is more important than do those who made column choices. Similarly, reasoners who made a row choice should also believe that D_2 is less important than do those who made a column choice. The importance ratings as a function of selection type (row vs. column) and anchoring feature (top university vs. parties) are presented in Figure 2. As predicted by such a relevance-driven search hypothesis, those who made a row choice when the anchoring feature was concerned with entry to a top university appeared more convinced that university was important and less convinced that parties was important than did those who made a column

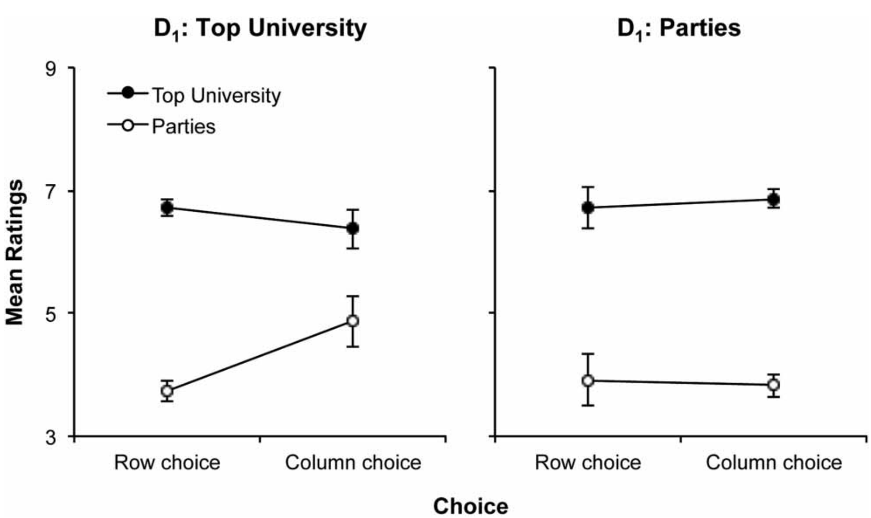

Figure 2. *Mean ratings of the importance of gaining entrance to a top university and the parties thrown annually as a function of selection type (row, column) and of the anchoring feature (D_1: top university or parties). Error bars represent 95% confidence intervals.*

choice. When the anchoring feature was the number of parties organized by the high schools, the pattern of ratings appeared to be reversed: Those comparing $P(D_1|H_X)$ and $P(D_1|H_Y)$ tended to believe that top university access was less important and parties was more important than those who made a column choice, although the differences seemed less pronounced. A 2 (anchoring feature) × 2 (feature importance) × 2 (choice) mixed ANOVA supported these observations. The main effect of the feature importance was significant, $F(1,\ 957) = 547.91$, $MSE = 3.04$, $p < .001$, $\eta_p^2 = .36$, as well as the main effect of choice, $F(1,\ 957) = 4.85$, $MSE = 2.39$, $p = .03$, $\eta_p^2 = .01$. The main effect of anchoring feature, however, was not significant, $F < 1$. The three-way interaction between choice, anchoring feature, and feature importance was significant, $F(1, 957) = 15.01$, $p < .001$, $\eta_p^2 = .02$. To further analyse this interaction, we conducted four follow-up *t* tests with Bonferroni corrections for multiple comparisons. These tests revealed that, when they were presented with the top university anchoring feature, participants who made a column choice did perceive parties as more important than those who made a row choice, $t(515) = -5.19$, $p < .001$. Although the means were in the expected direction, the difference in ratings for the importance of the top university feature did not reach statistical significance, $t(95.87) = 1.69$, $p = .10$ (corrected *df* for unequal variances). Neither of the corresponding comparisons when parties was the anchoring feature reached statistical significance, $t(89.19) = 0.31$, $p = .76$ for the importance of parties and $t(442) = -0.08$, $p = .43$ for the importance of getting access to a top university.

Discussion

The experiment reported here explored conditions under which so-called row selections— namely, comparing the probability of obtaining different data under one single hypothesis—would occur in "pseudodiagnosticity" tasks. The sample was large and more representative of the broader population than a traditional sample of university undergraduates. The data provided strong support for the hypothesis that the relevance of the anchoring feature encouraged individuals to make a row or column selection.

Gaining entrance to a top university was invariantly and significantly rated as more important than the number of parties organized by that school, across all four experimental conditions. In turn, determining which of the two alternatives proposed was best (viz. which high school to enrol in) was driven by a consideration of the high-relevance feature, not by a consideration of the low-relevance feature. Thus, in the two experimental conditions presenting the low-relevance feature as an anchor (viz., the number of parties organized yearly by the school), participants were still primarily interested in determining the proportion of students entering a top university. A large and significant majority opted for information on that latter feature, resulting in very high levels of column choices (about 72%), independently of the numerical value of this anchoring feature. Conversely, participants were strongly driven to consider the diagnosticity of the statistical value associated with the anchoring feature for multiple hypotheses when this feature was important. Specifically, when the percentage of students at high school X gaining entry to a top university was .20, 90% of the choices were row choices, a figure we believe is the highest rate of such choices in any reported experiments on "pseudodiagnostic" reasoning.

Note that while the order of the three choice types (row, column, diagonal) remained constant in all experimental conditions, a mindless "pick-the-first-option" strategy cannot explain the pattern of results given the effect size of the reversal in preferences (whereas 373 out of 516 participants chose the first option with the low-relevance anchoring feature, 440 out of 524 chose the second option with the high-relevance feature). Rather, the data suggest that the relative importance attributed to the two features presented strongly determined the choice preferences, not the order of the alternatives presented to participants.

Choice preferences elicited by the high-relevance anchoring feature also varied as a function of the initial numerical value presented. Contrary to Mynatt et al.'s (1993) conclusions, the initial value of $P(D_1|H_X)$ can impact data selection in action problems: When the high-relevance feature specified a high proportion of students enrolling in a top university, some of our participants became interested in determining the number of parties held at high school X. In this condition, the row choice was still the dominant response (78%) but the number of column choices more than doubled to 20%. In turn, when the relevance of the anchoring feature was low (e.g., the frequency of high-school parties), participants' choices were no longer affected by the initial value associated with the anchoring feature.

Dual-process theory

Our findings raise interesting issues for a dual-process theoretical account of information selection as proposed by Evans, Venn, and Feeney (2002) to explain patterns of information selection in inference problems. According to the heuristic–analytic theory of reasoning (Evans, 2006), reasoning tasks are initially processed heuristically through a preconscious pragmatic analysis, which will highlight relevant features and introduce relevant prior knowledge. Such heuristic processing is assumed to give rise to a selective representation, which may exclude logically relevant information or include logically irrelevant information. These representations may in turn highlight default responses, which may or may not be inhibited or altered in a second stage of processing, when the task is consciously and rationally analysed.

Akin to what we observed, Evans et al. (2002, Experiment 1) found that manipulating the perceived importance of the evidence had an effect on information selection. However, their findings showed that such a relevance manipulation was not sufficient to increase the prevalence of row selections. Instead, their manipulation resulted in an increasing number of participants switching their focus on both the alternative hypothesis and the alternative feature (the so-called "diagonal" selection) as the value of $P(D_1|H_X)$ decreased. The authors argued that the increase in diagonal selections reflected their participants' dismissal of H_X as the true hypothesis when the value $P(D_1|H_X)$ was lower than expected. As a result, participants favoured hypothesis H_Y, assumed that the $P(D_1|H_Y)$ value would be in line with their expectations, and sought to confirm that H_Y was the appropriate choice by switching to $P(D_2|H_Y)$.

The data obtained in the present experiment suggest that similar processes may take place with action problems. Indeed, a diagonal choice was about 10 times more likely when the relevance of the anchoring feature was low (i.e., describing the proportion of school parties) than when it was relevant (i.e., describing the admission rates to a top university), independently of the numerical value associated with the anchoring feature. This suggests that when the $P(D_1|H_X)$ anchor value was related to an irrelevant feature D_1, (some) participants also inferred that H_X was inappropriate and therefore sought information about the alternative H_Y on the relevant feature D_2 by making a switching selection.

As their manipulation of expectations did not lead to an increase in row data selection, Evans et al. (2002) concluded: "In accordance with the dual process theory, producing normatively rational diagnostic reasoning on the standard [pseudodiagnostic] task will require us to influence the explicit reasoning system" (p. 40). By contrast, Feeney, Evans, and Venn (2008) showed that simply offering anchoring statistical information about a rare feature was enough to increase the rate of row choices. Our results corroborate and extend Feeney et al.'s findings and therefore also contrast with Evans et al.'s (2002) assumption by demonstrating that, once the relevance constraint is satisfied, reasoners *can* engage in diagnostic reasoning comparing values associated with a unique feature across alternative hypotheses even in the absence of explicit instructions to consider the alternative option. Indeed, when D_1 was relevant, a large majority of participants chose to make row selections.

Unlike Evans et al. (2002) and Feeney et al. (2008), however, the high-school scenario used in our experiment was framed as an action problem rather than an inference problem. In addition, our manipulation of feature relevance called upon individuals' personal prior beliefs (i.e., the belief that admission rates to a top university are a more relevant feature than the frequency of parties for choosing a school). By contrast, Evans et al.'s

(2002, Experiment 1; see also Feeney et al., 2008, Experiment 3) manipulated expectations about the $P(D_i|H_j)$ values by telling their participants that both D_1 and D_2 features were important for the protagonist in their cover story. Thus, future research will need to establish whether manipulating the relative relevance of D_1 and D_2 through experimental manipulations will elicit high rates of row choices in action problems without explicit instructions to consider the alternative hypothesis.

Our data also suggest that individual perceptions of the relative relevance of both D_1 and D_2 may actually determine whether a row or a column choice will be preferred. Thus, when provided with information about success rates for gaining access to a top university, the few participants who made a column choice—and hence selected the $P(\text{parties}|H_X)$ information—also believed that parties were more important than did those who chose to inspect $P(\text{top uni}|H_Y)$. This was the only comparison that reached statistical level of significance, although relevance ratings were always in line with predictions based on a relevance-driven search account. This somewhat limited finding, however, could be attributed to the large effect of feature relevance on the patterns of choice, which may have limited the statistical power of the tests performed because of the resulting unequal cell sizes for the row and column choices. Future research could further explore this finding by using two features that differ in perceived relevance but where the consensus on which is the most relevant is not so clear-cut. For example, when deciding which lawn mower to purchase, men may be more interested in getting a sturdy machine whereas women might prefer a lightweight one. Presenting information about these two features should lead to reverse patterns of choice depending on participants' gender and anchoring feature.

Finally, a caveat. The two features employed in this study (rate of access to university and number of parties) were represented on very different scales. However, it is not clear how this difference might have influenced the results beyond the patterns observed and predicted from the relevance-driven information search hypothesis. Of course, some participants might have translated party frequencies as a rate over Fridays in a given focal set (a year, or a school year). Certainly, future research efforts may explore more systematically whether the scales that define the statistical information presented to participants influence choice preferences.

The normative status of choice preferences

The traditional normative analysis of choice selection (e.g., Doherty et al., 1979) in the pseudodiagnosticity paradigm has been the subject of much debate recently. Crupi et al. (2009) analysed information choices in terms of expected gains in epistemic utilities and showed that choosing the column value when the anchoring value is greater than 50% can be considered as a strategy that maximizes a Bayesian truth-seeking agent's epistemic utility, provided that some conditions apply. One of those conditions is that those rational agents apply Keynes' (1921) *principle of indifference*, which prescribes uniform distributions over sets of mutually exclusive and exhaustive events. Crupi et al. (2009) present this principle as "a celebrated epistemological principle . . . advocated in various guises by chief Bayesian theorists such as Keynes (1921)" (p. 975). This may have been true then, but this principle is no longer widely accepted as a normative benchmark. Rather, it is known to yield a number of unsolvable paradoxes with "alarming ease" (Howson & Urbach, 1993, p. 59). For instance, the "wine–water paradox" (Gillies, 2000, p. 38) shows how applying the principle of indifference leads to assigning different probabilities to the same events, depending on how they are described.

In fact, since the publication of the Ramsey–De Finetti theorem (Howson & Urbach, 1991), modern Bayesian theorists have overcome these inconsistencies by considering that the principle of indifference was simply unnecessary: Those theorists assume instead that there are no unique a priori "rational beliefs" (Gillies, 2000). Accordingly, and contrary to what Crupi et al. (2009) posited, rational agents' distributions of the hidden values in pseudodiagnostic tasks ought not to be uniformly distributed, as long as those probabilities are coherently distributed—that is, in a way that satisfies the mathematical calculus of probability.

Yet, once the condition for uniform distribution is forsaken, it is not clear that Crupi et al.'s (2009) account provides a definitive normative account of optimal information selections in so-called "pseudodiagnostic" tasks. In fact, Tweney et al. (2010) demonstrated that by simply assuming that unknown values belong to the [0, 1] interval, without positing any constraint on their distribution, the only value that will allow the computation of a precise estimate for $P(H_X)$ upon observing D_1 is the row value. The column value, contrastingly, will only allow one to compute a vague estimate, in the form of an interval of possible values for $P(H_X)$, which is larger than the interval one can compute based on the anchoring value alone. By this analysis, choosing to uncover the diagonal value leads to the vaguest estimate of all—namely, the conclusion that $P(H_X) \in [0, 1]$. These authors thus concluded that the only rational choice in those tasks was therefore to uncover the row value, as originally argued.

While we believe that these issues are important, the empirical contribution of the present study was not designed to address this debate: The data reported do not provide evidence to support the selection of a normative analysis rather than another to gauge the value of information selection, nor is our experiment predicated on one particular normative model. This paper focuses on the factors that determine choice preference: The relevance of the anchoring feature determined to a very substantial degree whether participants chose to learn more about the column value, $P(D_2|H_X)$, or the row value, $P(D_1|H_Y)$.

Conclusion: Rationality and motivated hypothesis testing

So-called "pseudodiagnostic" information selection is a phenomenon that has been well established. Nonetheless, in a number of experiments, diagnostic reasoning has also been observed, certainly in the action problems of Mynatt et al. (1993), but also in inference problems: For example, over 50% of the information selection choices in Evans et al.'s (2002) first experiment were row choices. Yet, such "optimal" behaviour—if one accepts Tweney et al.'s (2010) normative analysis of these tasks—is commonly overlooked or underemphasized. The focus on documenting errors in thinking shifts attention away from efforts aimed at understanding the inferential contexts where people make their information selection choices. Casting rationality in terms of an agent's motivation has been productively applied to work on the selection task (Evans & Over, 1996; Evans, Over, & Manktelow, 1993) and the 2–4–6 task (Vallée-Tourangeau & Payton, 2008). The present research suggests that stripping pragmatic considerations from scenarios used to study information selection may have created a misleading picture. Rather than being poor, irrational selectors of information, people adapt their information selection strategies to both their goal and their environment.

Original manuscript received 2 March 2009
Accepted revision received 14 July 2011
First published online 21 November 2011

REFERENCES

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

Crupi, V., Tentori, K., & Lombardi, L. (2009). Pseudodiagnosticity revisited. *Psychological Review*, *116,* 971–985.

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

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

Evans, J. St. B. T., & Over, D. E. (1996). Rationality in the selection task: Epistemic utility versus uncertainty reduction. *Psychological Review*, *103,* 356–363.

Evans, J. St. B. T., Over, D. E., & Manktelow, K. I. (1993). Reasoning, decision making and rationality. *Cognition*, *49,* 165–187.

Evans, J. St. B. T., Venn, S., & Feeney, A. (2002). Implicit and explicit processes in a hypothesis testing task. *British Journal of Psychology*, *93,* 31–46.

Feeney, A., Evans, J. St. B. T., & Venn, S. (2008). Rarity, pseudodiagnosticity and Bayesian reasoning. *Thinking and Reasoning*, *14,* 209–230.

Gillies, D. (2000). *Philosophical theories of probability*. London, UK: Routledge.
Howson, C., & Urbach, P. (1991). Bayesian reasoning in science. *Nature*, *350,* 371–374.
Howson, C., & Urbach, P. (1993). *Scientific reasoning: The Bayesian approach* (2nd ed.). Peru, IL: Open Court.
Kern, L., & Doherty, M. E. (1982). "Pseudodiagnosticity" in an idealized medical problem-solving environment. *Journal of Medical Education*, *57,* 100–104.
Keynes, J. M. (1921). *A treatise of probability*. London, UK: Macmillan.
Maggi, J., Butera, F., Legrenzi, P., & Mugny, G. (1998). Relevance of information and social influence in the pseudodiagnosticity bias. *Swiss Journal of Psychology*, *57,* 188–199.
Mynatt, C. R., Doherty, M. E., & Dragan, W. (1993). Information relevance, working memory, and the consideration of alternatives. *Quarterly Journal of Experimental Psychology*, *46A,* 759–778.
Savage, L. J. (1954). *The foundations of statistics* (2nd ed.). New York, NY: Dover Publications.
Trope, Y., & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology*, *43,* 22–34.
Trope, Y., & Bassok, M. (1983). Information-gathering strategies in hypothesis-testing. *Journal of Experimental Social Psychology*, *19,* 560–576.
Tweney, R. D., Doherty, M. E., & Kleiter, G. D. (2010). The pseudodiagnosticity trap: Should participants consider alternative hypotheses? *Thinking & Reasoning*, *16,* 332–345.
Vallée-Tourangeau, F., & Payton, T. (2008). Goal-driven hypothesis testing in a rule discovery task. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2122–2127). Austin, TX: Cognitive Science Society.

APPENDIX

The high-school scenario in the low-relevance anchor (number of parties) and low anchor value (one party a month) experimental condition read as follows (followed by the English translation):

Imaginez qu'un lycéen envisage de s'inscrire à une classe préparatoire aux grandes écoles. Il a le choix entre deux lycées: Le lycée X et le lycée Y. Il recherche un lycée qui offre des animations culturelles et sociales ainsi qu'une formation solide; à terme, il aimerait intégrer une grande école. Il s'est renseigné sur chacun des lycées et a remarqué que *le lycée X organise une soirée le dernier vendredi de chaque mois (et 10 au total durant l'année scolaire)*.

Imaginons qu'il puisse recueillir une seule autre information avant de faire son choix. A votre avis, de quelle information supplémentaire devrait-il prendre connaissance afin de décider dans quel lycée s'inscrire? (Veuillez cocher une seule case.)

- Le nombre de soirées organisées par le lycée Y chaque année.
- Le pourcentage des étudiants en classe préparatoire au lycée X qui ont réussi à intégrer une grande école.
- Le pourcentage des étudiants en classe préparatoire au lycée Y qui ont réussi à intégrer une grande école.

Imagine an adolescent who is choosing a prep course to prepare for the entrance exams of higher education establishments. He can choose from two senior high schools: high school X and high school Y. He is looking for a school that offers social and cultural activities as well as a solid education. Upon completion, he would like to gain entrance to a top university. He sought information about both high schools and noted that high school X organizes a party once a month on a Friday evening (and organizes 10 parties altogether during the school year).

Imagine that he can obtain only one other piece of information before making his choice. In your opinion, which additional piece of information should he select before deciding to enrol in one of the two high schools? (Only check one of the boxes.)

- The number of parties organized by high school Y during the school year.
- The percentage of students graduating from high school X who gained entrance to a top university.
- The percentage of students graduating from high school Y who gained entrance to top university.