Running head: BASE RATES IN CAUSAL LEARNING

When Learning Order Affects Sensitivity to Base Rates: Challenges for Theories of Causal Learning

Ulf-Dietrich Reips

Department of Psychology

University of Zürich, Switzerland

and

Michael R. Waldmann

Department of Psychology

University of Göttingen, Germany

Address for proofs:

Ulf-Dietrich Reips
Universität Zürich
Sozial- und Wirtschaftspsychologie
Binzmühlestrasse 14/ Box 13
8050 Zürich
Switzerland
E-mail: u.reips@psychologie.unizh.ch
Tel.: +41-44-63 44122
URL: http://www.psychologie.unizh.ch/sowi/reips/reipspers.html

Abstract

In three experiments we investigated whether two procedures of acquiring knowledge about the same causal structure, *predictive learning* (from causes to effects) versus *diagnostic learning* (from effects to causes), would lead to different base rate use in diagnostic judgments. Results showed that learners are capable of incorporating base rate information in their judgments regardless of the direction in which the causal structure is learned. However, this only holds true for relatively simple scenarios. When complexity was increased, base rates were only used after diagnostic learning, but were largely neglected after predictive learning. It could be shown that this asymmetry is not due to a failure of encoding base rates in predictive learning because participants in all conditions were fairly good at reporting them. The findings present challenges for all theories of causal learning.

Acquiring causal knowledge about the world belongs to our most important competencies. Causal knowledge allows us to predict future events or explain the occurrence of present events. Causal information may be presented in different ways (see Reips, 1998; Waldmann, 1996, 2000, 2001; Waldmann & Holyoak, 1992; Waldmann, Holyoak, & Fratianne, 1995)(see Figure 1). Medical textbooks frequently are structured according to disease categories that are the causes of symptom patterns. Students of these textbooks will learn to predict symptoms on the basis of diseases. Thus, they acquire *predictive* knowledge (i.e., from causes to effects).

However, in other contexts information may be presented in the *diagnostic* effect-cause direction. A physician who sees a patient for the first time will check the symptoms (i.e., effects) and then attempt to settle on a diagnosis of the probable disease (i.e., the cause). Also, textbooks that focus on differential diagnosis typically present causal knowledge in the diagnostic direction.

When a physician knows that a patient has a specific disease, she can predict the future symptoms. These predictions require knowledge about the strength of the causal relations between disease and symptoms but it does not matter whether the disease is rare or frequent. In contrast, diagnoses of likely diseases from observed symptoms are only appropriate if both are taken into account, causal strength and the frequencies (i.e., base rates) of possible diseases. Thus, due to this asymmetry it may well be that predictive but not diagnostic learning leads to a neglect of base rates which may entail erroneous judgments if in the future the predictive learners will be asked to make diagnostic judgments.

Very little is known about the relationship between these two different types of learning contexts, predictive and diagnostic learning, and their influence on the resulting representation of causal knowledge. The main goal of the present research is to close parts of this research gap by focusing on one important aspect of normative diagnostic reasoning, *sensitivity to the base rates* of the causes of the observed pattern of symptoms.

The Use of Base Rates in Judgments

One of the most discussed findings in the literature on judgment and decision making is Kahneman and Tversky's (1973) discovery of base rate neglect. In several experiments they have found that people do not adequately take into account the base rates of events in diagnostic tasks. An example from the medical domain was investigated by Gigerenzer and Hoffrage (1995). In one of their experiments they told participants that the probability of breast cancer is one percent. Moreover, if the patient had breast cancer, it would be detected in a mammography in 80 percent of the cases. If the patient did not have breast cancer, it would be incorrectly detected in a mammography in 9.7 percent of the cases. Presented with this information, participants tended to drastically overestimate the likelihood of breast cancer for a particular woman with a positive mammography. Students and professional physicians gave assessments above 50 percent (Eddy, 1982), whereas the actual conditional probability is only 7.8 percent. These results indicate that people neglected or underused the information about the base rates and tended to base their judgments on the information about the likelihoods of positive tests.

However, the stability of the base rate neglect phenomenon has been called into question (see Koehler, 1996, for a review). Studies that conveyed the information in summary format have shown that subtle variations of the wording of the task may affect the degree of appreciation of base rates. Tversky and Kahneman (1980) had already shown that base rates tend to be used more frequently when they are causally motivated. Others have also demonstrated that the perceived relevance of base rates plays an important role (e.g., Ajzen, 1977; Bar-Hillel, 1980; Gigerenzer, Hell, & Blank, 1988). Moreover, Gigerenzer and Hoffrage (1995) have shown that base rates are used more often when the information is given in frequency format, at least in situations in which all presented frequencies are related to an identical and clearly defined sample (Fiedler, Brinkmann, Betsch, & Wild, 2000).

Not all causal knowledge is acquired on the basis of summary and frequency formats. When base rate information is conveyed directly through trial-by-trial experience people can fully make use of it (e.g., Christensen-Szalanski & Beach, 1982; Gluck & Bower, 1988; Shanks, 1990; Spalding & Murphy, 1999). However, some studies have found base rate neglect or even an inverse base rate effect with trial-by-trial learning procedures (e.g., Gluck & Bower, 1988; Medin & Edelson, 1988; Kruschke, 1996). Holyoak and Spellman (1993) have suggested that base rates are implicitly used during learning but may be neglected when the test question requires explicit use of base rates. Although this factor seems to be important, there are also studies showing base rate neglect with directly experienced data and more implicit tests (Goodie & Fantino, 1995; Lovett & Schunn, 1999). The present research extends this research by investigating further factors that might affect base rate sensitivity in trial-by-trial learning tasks.

Predictive vs. Diagnostic Learning

Medical diagnosis is arguably the most important domain in which base rates should be used. Physicians or medical advisors who neglect base rates may give erroneous advice or initiate inadequate treatments. The goal of our studies is to focus on trial-by-trial learning in a causal task from the medical domain, and investigate the conditions under which base rates are used. We are going to study predictive and diagnostic learning of the same causal structure of diseases and symptoms, and study their impact on the use of base rates.

In order to understand our experimental paradigm (see Figs. 1 and 2), it is important to note the distinction between *temporal order* (cue and outcome) and *causal order* (cause and effect). Cues and outcomes are generic terms for antecedent and consequent events regardless of their causal description. A cue can represent a cause and an outcome an effect (predictive task) or a cue can represent an effect and an outcome a cause (diagnostic task).

In the *diagnostic learning task*, participants will be given information about symptoms (i.e. the effects of diseases) as cues and will be asked to diagnose the disease (i.e., the cause of

the symptoms). Each disease has a unique symptom that is only caused by this disease (e.g. stiff joints caused by the disease ”pleroia” in Fig. 2), and an ambiguous symptom that is caused by two competing diseases (e.g. shortness of breath is present if a patient has pleroia or spetitis). The base rates of these two diseases vary. To test sensitivity to base rates, the final test questions require diagnostic inferences in the effect-cause direction. The crucial test involves the ambiguous symptoms (e.g., shortness of breath in the example in Fig. 2). Since each of these symptoms is deterministically caused by two competing diseases (e.g., both pleroia and spetitis always lead to shortness of breath), the diagnostic judgments should reflect their base rates. Given the ambiguous symptom as a single cue in the test phase, the more frequent disease should be judged as more likely than the less frequent disease. In the example, given only the information that a patient has shortness of breath and one of the two diseases, a good diagnostician should assume that the patient has pleroia with a likelihood of 75%.

In the *predictive* version of the task we will present the same diseases with identical causal structures and base rates in the predictive direction from the causes (the diseases) to their effects (symptoms)(see Fig. 1). Thus, learners will be given the individual diseases as cues and will have to predict the two symptoms that are deterministically caused by the diseases. For example, given pleroia participants will learn to predict the presence of stiff joints and shortness of breath in patients. In the test phase, symptoms will be given as cues (as in the diagnostic condition), and learners will be asked to assess the likelihood of the diseases. Thus, in both conditions participants learn the same causal structures with each disease causing two symptoms and with diseases varying in frequency. Moreover, in both tasks participants are requested to make diagnostic judgments based on individual symptoms as cues. The only difference is that in the diagnostic learning condition participants receive symptoms as cues and learn to diagnose the associated diseases, whereas in the predictive learning condition they receive diseases as cues and learn to predict the associated symptoms.

Normatively, base rates should be used regardless of learning direction. Also if the use of base rates hinges on the presentation of trial-by-trial information (Christensen-Szalanski & Beach, 1982) or on the presentation of natural frequencies that are related to identical samples (Fiedler et al., 2000; Gigerenzer & Hoffrage, 1995), sensitivity to base rates should be equal in both conditions. In both learning procedures participants are presented with information of sequences of individual patients, and are asked to give identical diagnostic judgments.

The comparison between predictive and diagnostic learning in the present experiments differs from the tasks used in previous studies designed to test causal-model theory (e.g., Waldmann & Holyoak, 1992; Waldmann, 2000, 2001). In their experiments the cues were either characterized as effects (diagnostic learning) or causes (predictive learning) of the outcomes, thus varying causal models of the same structures while keeping the learning order and the test questions constant. The goal of the present studies is to keep the causal model of a structure constant while studying the impact of learning order (see Figs. 1, 2)(see Cobos, López, Caño, Alvarez, & Shanks, 2002; Yamauchi & Markman, 1998; Yamauchi, Love, & Markman, 2002, for related paradigms).

Competing Theories of Learning

During the last few years there has been a debate between associative and cognitive accounts of causal learning (see De Houwer & Beckers, 2002; De Houwer, Beckers, & Vandorpe, 2005). Thus, it is interesting to consider the predictions of these theories for our tasks.

Causal-Model Theory. We will discuss causal-model theory as a representative example of rational models of complex causal model learning because it is the only theory that has addressed the differences between predictive and diagnostic learning so far (but see Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004, for a related theory). Causal-model theory assumes that learners form a representation of causal models regardless of the order in which learning information is presented. These causal models contain information about how causes are related to effects (i.e., causal structure) along with estimates about the parameters (e.g.,

causal strength, base rates) that are gleaned from the learning data. It is typically assumed that the parameters are estimated on the basis of the observed frequencies in the learning data (Waldmann & Holyoak, 1992; Waldmann & Hagmayer, 2005).

The focus of the present experiments is the use of base rates in diagnostic judgments. The predictive learning condition is the most interesting condition for testing causal-model theory because in this condition the task at test is directed in the opposite direction to learning order. In this condition participants are required to switch from predictive learning to diagnostic judgments. Only if learners correctly acquire a causal model and its parameters (causal strength, base rates), and if they can correctly access the model in both predictive and diagnostic directions, normative judgments are to be expected for the diagnostic test questions after predictive learning. Causal-model theory predicts that learners should attempt to acquire causal-model knowledge regardless of learning order, and hence be sensitive to base rates in both learning conditions, predictive and diagnostic learning. In the General Discussion an extended version of causal-model theory is discussed that is sensitive to the complexity of the task (see also Reips, 1998).

Associative Theories. We will discuss the Rescorla-Wagner theory of associative learning as a representative example of this class of theories. This theory has been applied to model sensitivity to base rates (Gluck & Bower, 1988; Shanks, 1990). Associative theories would model the diagnostic task with symptoms as cues and diseases as outcomes. Since the underlying associative learning rule attempts to reduce errors, eventually the learning model would correctly diagnose the diseases. These diagnoses would be sensitive to the base rates. It is interesting to note that these models explain sensitivity to base rates without having to separately represent base rates. The diagnoses are simply a consequence of combining associative weights that are tuned to take the objective base rates into account.

In the predictive context, associative models would assign the diseases to the cue layer and the symptoms to the outcome layer. In this context the models attempt to correctly predict

the outcomes (i.e., the symptoms). At the learning asymptote the model should have learned maximal weights representing the deterministic relations between each disease and its two associated symptoms. Since outcomes, according to the Rescorla-Wagner and related theories, do not compete with respect to their cues, the associative weights are solely dependent on the contingencies between diseases and symptoms. Base rates do not affect the outcome of learning, at least at the asymptotic stage.

What would these models predict for the diagnostic judgments? Since our test phase requires assessments opposite to the learning direction (from outcomes to cues), additional assumptions need to be made. A simple assumption would be that people use the associative weights from the learning task also in this phase. In this case associative theories would predict base rate sensitivity after diagnostic but not after predictive learning. Possible extensions of this basic account and alternative theories will be discussed in the General Discussion.

One interesting empirical question will be whether participants encode the frequencies of the causes even when they are ignored in the actual diagnoses. In the predictive learning context the task does not require an encoding of frequencies. However, it has been argued that frequencies often are encoded automatically (Hasher & Zacks, 1979; Reber, 1993). Whereas causal-model theory assumes that people use conditional frequency information to arrive at inferences (Waldmann & Holyoak, 1992), the Rescorla-Wagner model does not predict storage of frequency information. It may also happen that people store the base rates but do not use them in their judgments. If that was the case, an underuse of base rates would support the notion that base rate neglect in our task is not a consequence of a failure to encode frequency information during learning but rather a consequence of the fact that participants did not embody base rates in the trial-by-trial judgment procedures. Frequencies may be stored passively but still not used in the diagnoses. This finding would place further constraints on theories of causal learning.

Experiment 1

The first experiment is designed to test whether learners have the competency of correctly incorporating base rate information regardless of the sequence of learning. This competency is predicted by causal-model theory, but would be at odds with associative theories. To test this hypothesis, participants in the predictive and the diagnostic learning conditions received identical learning materials that contained information about two different diseases (causes) and their three associated symptoms (effects). The only difference in the learning phase was that participants in the predictive condition received information about the diseases as cues and had to learn to predict symptoms as outcomes, whereas in the diagnostic condition learners received information about the symptoms first as cues and had to learn to diagnose the diseases. After the learning phase all participants were requested to give assessments of the probability of the diseases given information about the presence of individual symptoms. Accordingly, the test phase was directed in the diagnostic direction from effect cues to their causes.

To test whether participants were sensitive to base rates we used a causal structure ("M-structure") in which each disease had two symptoms, one of which was unique for the disease. The other symptom was shared with a second disease and thus ambiguous (see Fig. 2). All symptoms were deterministically caused by the associated diseases. We varied the base rates of the two diseases that competed for the explanation of the ambiguous symptom (see Medin & Edelson, 1988, for a similar task). Base rate sensitivity implies that participants would give the more frequent disease a higher probability than the rare disease when the ambiguous symptom is present.

Participants and Design

There were 24 participants, all students from the University of Tübingen, who received either participation credit or DM 5. Half of this group was randomly assigned to either of the two learning conditions, predictive or diagnostic learning.

Material

Instructions and learning trials in all experiments were presented in German. As symptoms we used stiff joints, shortness of breath, and muscle cramps. The fictitious diseases were "pleroia" (frequent) and "spetitis" (rare). The role of each symptom as either a unique or an ambiguous cues was counterbalanced. The order of trials was randomized within blocks, with each block representing a complete M structure.

The causal "M-structure" underlying the learning material was constructed as follows (see Fig. 2): There are two diseases and three symptoms. Each disease deterministically causes two symptoms. One of these two symptoms is *ambiguous* in that it is an effect of either diseases, while the other two *symptoms* are each caused by one disease only (i.e., unique symptoms). For example, the disease pleroia causes both stiff joints and shortness of breath, and the disease spetitis causes the symptoms shortness of breath and muscle cramps. Thus, shortness of breath is the ambiguous symptom because it does not allow to decide between the diagnoses pleroia and spetitis. Base rates were manipulated in a 3:1 ratio, meaning that one disease within the M-structure was three times as frequent as the other disease. In the experiment participants saw 24 times the frequent disease and its associated symptoms on individual index cards, and 8 times the rare disease and its associated symptoms.

Dependent measures

Our data analysis in all experiments focuses on diagnostic decisions based on the ambiguous symptoms because only these are indicators of base rate use. Base rate sensitivity implies that the more frequent disease is seen as a more probable cause than the rare disease when an ambiguous symptom is present and no other information is available. Sensitivity to base rates is indicated when participants give a higher probability rating for the frequent disease than for the rare disease when confronted with the ambiguous symptom that is deterministically caused by either disease. Thus, we generally defined base rate use as the difference between

participants' ratings or proportion of choices of the frequent disease minus their ratings or proportion of choices of the rare disease. If applicable, the differences were averaged over all causal structures (see Experiments 2 and 3), resulting in the *measure of base rate use*. If the measure takes a positive value, it indicates base rate use, if it is zero it indicates no base rate use. Negative values would show a reversed use of base rates.

Procedure

Participants were run individually. Before going through the learning trials participants received typed instructions (in German). To facilitate thorough reading of the instructions all participants were told that they would be asked to summarize the written instructions once having read through them. Participants were asked by the experimenter to re-read the instructions whenever their oral summary indicated a misunderstanding of the instructions. In the instructions, all participants were asked to imagine being a guest in a special clinic for viro-neuronal tropical diseases for one day. Participants in the *diagnostic learning condition* were told that they would be learning to diagnose diseases, and that their task was to diagnose patients' diseases based on information about the symptoms these patients exhibited. Participants in the *predictive learning condition* were told that they were going to learn to predict symptoms, and that their task was to predict patients' symptoms on the basis of information about the disease on the patients' cards. After summarizing the instructions, participants began with the learning task. Descriptions of patients were presented by the experimenter, one by one, on 32 index cards. Each card displayed two symptoms on one side and one disease on the other side. The cards were presented in blocks of four trials, each block containing all trials for the base rate distribution within the M-structure. Card order was randomized within each block. The participants in the *diagnostic learning condition* were shown the side with the symptoms first. After having announced the diagnosis, learners were shown the back of the card, which showed the patient's disease. The participants in the *predictive learning condition* were presented the side with the disease first. After having

announced the prediction they were shown the back of the card, which showed the patient's symptoms. Figure 3 shows a schematic description of the procedure used.

After the learning phase, participants were handed typed sheets with rating instructions and rating scales. In the instructions participants were told to imagine receiving information about the next patient arriving at the clinic. Then it was pointed out that in this phase participants would only receive information about a single symptom of the patient. The task was to rate the probability of the patient having the respective disease on a scale ranging from very improbable (0) to very probable (100). Thus, participants were asked to give ratings of the diagnostic relation between individual symptoms and the diseases.[1] In addition to the ratings, we asked participants to make a forced choice between the diseases in the presence of each of the symptoms. We also asked participants to give relative frequency estimates for the diseases on a scale from 0 to 100 percent.

Results and Discussion

We conducted an ANOVA with the average differences between probability ratings for the two diseases in the ambiguous symptoms' presence. The explicit measure of base rate use was not significantly different for causal learning direction, $F(1, 22)=0.41$, $MSE=1023.49$, n.s. (see Fig. 4). The means of the measure of base rate use were 18 ($SD=29$) in the diagnostic learning condition and 26 ($SD=35$) in the predictive learning condition. In both learning conditions there was clear evidence for base rate use, meaning that the mean probability ratings for the more frequent disease were higher than for the less frequent disease.

In sum, Experiment 1 shows roughly equal amounts of base rate appreciation after predictive and diagnostic learning (see Figure 4). In the diagnostic learning condition there were four and in the predictive learning condition there were five participants who made use of the base rate information. No one gave a higher rating for the rare as compared to the frequent disease. The analysis of the forced choice data also revealed a complete lack of an asymmetry between the conditions. The same number (9) of participants in both causal

conditions chose the frequent disease when confronted with a patient showing the ambiguous symptom.

The ratings for frequencies of the diseases also turned out to be similar in both causal conditions (see Fig. 5). An analysis of variance on the differences between the average frequency ratings for the frequent diseases versus the rare diseases with the factor causal learning direction as the independent variable showed no statistically significant difference, $F(1, 22)=0.01$, $MSE=448.11$, n.s. The average means for differences were 48 ($SD=24$) in the diagnostic learning condition and 49 ($SD=17$) in the predictive learning condition. The main effect between the learning conditions was not significant. Thus, participants appeared to encode frequency information regardless of learning direction.

In the present experiment participants tended to be equally sensitive to base rates in both the diagnostic and the predictive learning conditions even though the predictive learning task does not require this sensitivity to achieve error free performance. The results are consistent with the predictions of causal-model theory but are at odds with the Rescorla-Wagner theory and related models. Further evidence for a non-associative account is the fact that learners were aware of the different base rates of the diseases in their frequency judgments.

The results of the present experiment are consistent with our previous research supporting causal-model theory, which showed that learners try to correctly represent causal knowledge regardless of the sequence of the learning input (Waldmann & Holyoak, 1992; Waldmann, 1996, 2000, 2001). Whereas previous experiments have demonstrated this skill in tasks in which cues and outcomes were kept constant while varying the underlying causal model, the current experiment provides the first evidence for the competency to correctly learn about identical causal models irrespective of the sequence in which the elements of the models are experienced.

Experiment 2

Experiment 1 demonstrated sensitivity to base rates irrespective of learning order. This finding is consistent with causal-model theory, which claims that people attempt to form adequate representations of causal models regardless of the order in which knowledge is acquired (see Waldmann, 1996; Lagnado, Waldmann, Hagmayer, & Sloman, in press). With Experiments 2 and 3 we pursued the goal to investigate the boundary conditions of this competency. Previous research has shown that the competency to acquire knowledge about causal models can break down when complexity of the domain or the task is increased (De Houwer & Beckers, 2003; Reips, 1998; Waldmann & Walker, 2005). We therefore increased the complexity of the task by increasing the number of diseases and symptoms.[2] In Experiment 2 we presented a task with six diseases and nine symptoms (triple M-structure). As in Experiment 1, one of the two diseases within each M-structure and its symptoms was presented three times as frequent as the other disease within that structure.

We generally expected that the more difficult condition in which learning order and test order mismatch (predictive learning) should be particularly prone to performance deficits. Adequate learning of causal models requires the acquisition of knowledge of the structure and of the size of the parameters. A plausible strategy used by learners under taxing conditions might be to abandon the goal to form complete causal model representations that can be flexibly accessed, and fall back on learning only the information that is necessary to minimize errors in the current task (see also Lovett and Schunn, 1999). Reducing errors in diagnostic learning requires the diagnoses to be tuned to base rates; therefore we expected sensitivity to base rates in this condition. In contrast, predictive learning does not require sensitivity to base rates. In this condition, learners predict symptoms on the basis of more or less frequent diseases. Whereas diseases compete for explaining symptoms, there is no competition between symptoms that would require learners to take into account base rate information. Thus, learners in the predictive condition may correctly acquire the knowledge about the

structure of the causal models (i.e., M-structures with deterministic relations) but reduce the learning effort by ignoring parameters, such as the unequal base rates, that are currently not relevant for successful performance.

An interesting empirical question will be whether participants encode the frequencies of the diseases even when they are ignored in the diagnostic estimates. Whereas standard associative theories would not predict an encoding of frequency information, probabilistic theories (including causal-model theory) assume storage of frequencies. The potential dissociation between storage and use of base rates after predictive learning would place important constraints on theories.

Method

Participants and Design. There were 32 participants, mostly students from the University of Tübingen, who were recruited in the university cafeteria. They either received participation credit or were paid DM 8. Participants were randomly assigned to either the diagnostic learning condition or the predictive learning condition.

Material and Procedure. The most important difference to Experiment 1 was that we presented *three* M-structures instead of one, with six diseases (“terrigitis”, “spetitis”, “rutix”, “pleroia”, “bilea”, “althrax”) and nine symptoms (irritant cough, ear-ache, muscle cramps, hot flushes, skin rash, pain in the limbs, stiff joints, eye irritation, shortness of breath). As in Experiment 1 each disease deterministically caused one unique and one ambiguous symptom. The relative base rates of the two competing diseases were the same as in Experiment 1 (3:1); the names of diseases were randomly assigned to the base rates. Moreover, the nine symptoms were randomly assigned to the six diseases.

The learning trials were presented on a computer monitor using the Micro Experimental Laboratory (MEL) software. Except for the different frequency judgment scales (see below), participants received the same instructions and rating scales as in the first experiment, with an additional instruction on how to use the computer, and two new types of

questionnaires. Participants were informed that they would receive patient information on the computer. Each new display on the computer screen would represent one patient who had just been hospitalized. Then participants were instructed that the speed of the experiment would be self-paced: The experimenter would press the button to display the label of the disease and the symptoms of a patient only after the participant's verbal answer to a trial. The sequence of information was similar to the procedure in Experiment 1, which means that disease information was presented first in the predictive learning condition and symptom information first in the diagnostic learning condition. The next patient's information would be displayed only after the participant had studied the feedback and had said “ok” or “continue” or similar. In order to reduce task difficulty two sheets of paper that listed all possible diseases and symptoms were on display throughout the experiment. We set a learning criterion of two completely correctly answered blocks with a minimum of 48 trials and a maximum of 192 trials (which nobody reached). In addition to the ratings, we asked participants to make a forced choice between the diseases in the presence of each of the symptoms. The order of symptoms on the response sheets was randomized. Furthermore, we asked participants how frequent the diseases were, on a rating scale with the endpoints 1 (“very rare”) and 7 (“very frequent”).

Results and Discussion

The average means of ratings of the probability of the high frequency diseases conditional upon the ambiguous symptom were similar in the causal conditions. They were 63% (*SD*=25) after diagnostic learning and 61% (*SD*=20) after predictive learning. For the low base rate diseases, however, participants in the diagnostic condition rated the probability on average at 36% (*SD*=18), while the respective value in the predictive learning condition turned out to be 64% (*SD*=23)(see Fig. 6).

As in Experiment 1, we conducted an analysis of variance on the *measure of base rate use*, that is the difference between probability ratings for the two diseases in the ambiguous

symptom's presence, averaged over all M-structures. In the present experiment with the more complex causal structure the measure of base rate use was significantly different in the two contrasting conditions, $F(1, 30)=11.93$, $MSE=612.08$, $p<.01$. Base rate appreciation was higher after diagnostic learning ($M=27$, $SD=29$) than after predictive learning ($M=-3$, $SD=19$) following the learning of the triple M-structure. Only four out of 16 participants in the diagnostic learning condition gave ratings inconsistent with base rate use (three equal, one higher for the infrequent cause), in contrast to 13 out of 16 in the predictive learning condition (eight equal, five higher for the infrequent cause).

In contrast to the probability estimates, the analysis of the forced choice data revealed no asymmetry between the conditions. A Kruskal-Wallis test on the differences between the number of choices of the frequent diseases and of the rare diseases showed no statistically significant difference for causal learning direction, $\chi^2(1, N=31)=0.34$, n.s. (one participant did not fill out the choice questionnaire). Probably the choice measure is less sensitive to differences in the strength of base rate sensitivity than the probability measure because people might choose the more frequent disease even when they believe that the probabilities of the frequent and the rare disease are very close.

Ratings for frequencies of diseases again turned out to be similarly accurate in both causal conditions (see Fig. 7). An analysis of variance on the differences between the average frequency ratings for the frequent diseases versus the rare diseases with the factor causal learning direction as the independent variable showed no statistically significant difference, $F(1, 30)=0.47$, $MSE=1.26$, n.s. The average means for differences were 2.7 ($SD=1.2$) in the diagnostic learning condition and 3.0 ($SD=1.0$) in the predictive learning condition. There were no significant differences between learning conditions. These results suggest that participants acquired the base rates of the diseases in all conditions but used them differently in probability ratings depending on the learning condition.

In summary, in Experiment 2 we used a triple M-structure as the learning material. In contrast to the single M-structure of Experiment 1, there was a difference in base rate use after diagnostic versus predictive learning. These results contradict the predictions of causal-model theory, and support associative theories. A modified version of causal-model theory that assumes that learners may neglect information that does not seem crucial for successful performance may also account for the results (see General Discussion).

An interesting result concerns the direct assessments of frequencies. Although frequencies were neglected in the diagnostic probability ratings after predictive learning, participants still encoded the frequencies of rare and frequent diseases equally well in both learning conditions. Thus, the found asymmetries of base rate use are not a result of a failure of encoding base rates after predictive learning. Since basic associative theories do not predict the encoding of frequencies, this finding requires more complex models. This finding is also critical for causal-model theory which anticipates storage of frequency information but does not predict that encoded frequency information may not be used in diagnostic inference tasks (see General Discussion).

Experiment 3

The results of Experiment 2 suggest that the asymmetry of base rate appreciation is particularly strong when probability assessments were requested, and less strong with the choice measure. A plausible explanation of this possible difference may be that the choice measure is sensitive to small differences and therefore does not differentiate between different sizes of sensitivity to base rates. One goal of Experiment 3 was to use a more sensitive measure of diagnostic inference that is based on choice, but is still better comparable with the probability measure. Our focus on choice was motivated by our goal to use a more implicit measure of base rate sensitivity than explicit frequency estimates. In general, in the present experiment we were interested in exploring whether our findings are restricted to explicit measures or can also be replicated with more implicit measures. When looking for a more

sensitive implicit task that is not based on explicit frequency estimates we tried to capitalize on the finding that learners tend to *match probabilities* when making blocks of diagnostic decisions (see Reber, 1993). Therefore, in the test phase all participants received several blocks of individual diagnostic trials without feedback. Our goal was to use the relative frequencies of the diagnoses of the frequent and the rare diseases given the ambiguous symptoms as an implicit indicator of participants' probability estimates. Thus, different probabilities for the frequent and the rare diseases were taken as an implicit indicator of base rate sensitivity.

Participants and Design

There were 32 participants in this experiment who were randomly assigned to either the predictive or the diagnostic learning condition. They received either participation credit or DM 10.

Method and Procedure

The procedure remained largely unchanged from Experiment 2. Instead of the choice measure we used the new implicit probability measure. Again, participants sat in front of a computer screen, and were presented screen by screen with information about fictitious patients who had supposedly just been hospitalized. The procedure was self paced. Participants were either instructed to diagnose diseases or, in the predictive learning condition, to predict symptoms. Lists of symptoms and diseases were available throughout the experiment. In the learning phase we gave participants a minimum of four and a maximum of 16 blocks of 12 learning trials each, using a learning criterion of two completely correct blocks, in which they were presented with individual symptoms. In the test phase, participants saw single symptoms, and were requested to choose among the possible diseases as probable causes of the symptom. There were 48 test trials, presented in random order that presented individual symptoms. The frequencies of the individual symptoms mirrored two presentations of each of the three M-structures. Thus, each ambiguous symptom was presented eight times, each frequent unique

symptom was presented six times, and each rare unique symptom was presented twice. Participants received no feedback in the test phase.

The general procedure used in this experiment can be summarized as follows: (1) written general instruction, (2) computer instruction, (3) learning phase with feedback, (4) instruction for the test phase, (5) test phase of diagnostic judgments without feedback, (6) rating questionnaire, (7) frequency questionnaire. The same questionnaires were used as in Experiment 2.

Results and Discussion

One participant did not meet the learning criterion of two completely correct blocks so that this participant's data were not included in the following statistical analyses. Generally, the results were similar to those from Experiment 2. The average means of the ratings of the probability of the high base rate diseases were 73% (*SD*=22) in the diagnostic learning condition and 62% (*SD*=22) in the predictive learning condition. For the low base rate diseases, participants in the diagnostic condition rated the probability on average at 46% (*SD*=31), while the respective value in the predictive learning condition turned out to be at 58% (*SD*=13). On average (over all three M-structures), only six of the 16 participants in the diagnostic learning condition did not use base rates, in contrast to 12 out of 16 in the predictive learning condition.

As in the previous experiments, we conducted an ANOVA with the average difference between probability ratings for the two diseases in the ambiguous symptoms' presence. Again, this explicit measure of base rate use yielded significant results, $F(1, 29)=5.68$, $MSE=740.28$, $p<.05$. This finding replicates the results of the previous experiment for the explicit measure of base rate use.

The novel question in the present study was whether the asymmetry of base rate use would also be found in a more implicit measure that was closer to the learning task. Figure 8 shows the mean percentages of choices of the frequent diseases in the presence of the

ambiguous symptom: the high frequency disease was diagnosed in 70% of the trials in the diagnostic learning condition and in 58% in the predictive learning condition, whereas the low frequency disease was diagnosed in 20% in the diagnostic learning condition but in 41% in the predictive learning condition. Consequently, an analysis of variance with the average differences between percentages of choices for the two diseases in the ambiguous symptoms' presence (i.e., *implicit measure of base rate use*) again revealed a significant difference between the causal conditions, $F(1, 29)=5.96$, $MSE=1378.44$, $p<.05$.

Ratings of the frequencies of diseases once more turned out to be similar in both learning conditions, indicating no difference in the *encoding* of base rates. An analysis of variance on the differences between the average frequency ratings for the frequent diseases versus the rare diseases with the factor causal learning direction as the independent variable showed no statistically significant effect, $F(1, 29)=1.46$, $MSE=1.15$, n.s. The average means for differences were 3.4 ($SD=1.1$) in the diagnostic learning condition and 3.0 ($SD=1.1$) in the predictive learning condition. Again there were no significant differences between learning conditions. Figure 9 shows the average frequency ratings for the frequent and for the rare diseases separated by learning direction.

In summary, the results from this experiment confirm and expand our earlier findings. As in Experiment 2 we used a triple M-structure as learning material. Again, there was a pronounced difference in base rate use after diagnostic but not after predictive learning in the diagnostic ratings. The interesting novel question whether this effect would also show up with a more implicit measure of base rate was answered as well. While we observed a certain use of the base rates in the predictive learning condition in the patterns of choices (see Fig. 8) we were also able to replicate the basic asymmetry between the causal learning conditions in complex tasks with the new measure. The use of base rates in the implicit measure is clearly less pronounced in the predictive learning condition than in the diagnostic learning condition. We also replicated the finding of Experiments 1 and 2 that participants encoded the

frequencies of the diseases fairly well in both learning conditions. This shows again that knowledge of base rates is not sufficient for using them in diagnostic judgments.

General Discussion

The three experiments present challenges for a unified account of causal learning. Experiment 1 demonstrates that people incorporate base rate information in their diagnostic inference independent of the learning sequence in which the causal model was learned. However, when the complexity of the learning domain was increased in Experiments 2 and 3, performance deteriorated, especially when there was a mismatch between the learning sequence and the test sequence. Whereas learners were sensitive to the base rates of the causes in diagnostic test questions after diagnostic learning, they tended to neglect them after predictive learning. This asymmetry could be demonstrated in a probability rating task that required some abstraction from the learning task but was also prominent in a more implicit task in which participants matched the probability of the diseases in blocks of diagnostic judgments. Furthermore, it could be shown that this asymmetry is not due to a failure of encoding base rates in predictive learning. Participants generally remembered the frequency of the diseases fairly well in both learning conditions. In our view, these findings have theoretical as well as practical consequences.

Theoretical Challenges

The pattern of results in our experiments presents interesting challenges to extant theories. The competence of learners, displayed in Experiment 1, supports causal-model theory but presents problems for theories that model learning as solely directed from cues to outcomes (e.g., associative theories). These theories can explain base rate sensitivity after diagnostic learning, but, without an extension, are ill suited to model base rate sensitivity when learning proceeded in a direction opposite to the direction of the test questions (predictive learning and diagnostic testing).

One possible explanation of the results of Experiment 1 is to retain the basic version of the Rescorla-Wagner theory but make the assumption that learning was pre-asymptotic. As pre-asymptotic weights should be smaller for the rare diseases than the frequent diseases this may explain base rate sensitivity in Experiment 1 even after predictive learning. However, a shortcoming of this approach is that it would also erroneously predict base rate sensitivity in predictive learning in Experiments 2 and 3. Moreover, we used a learning criterion in these experiments and fairly simple deterministic structures so that this account seems implausible.

A possible extension of the basic associative theory would be to propose a model that learns bidirectional links between cues and outcomes (Shanks & Lopez, 1996). Such a model could propose that people simultaneously learn in both directions which would predict base rate sensitivity in both tasks. Alternatively one could propose that learners acquire associations between outcomes (i.e., symptoms), or between outcomes and the context. The first proposal would lead to stronger associations between the ambiguous symptom and the frequent symptom rather than the rare symptom, the second to a stronger association between context and the frequent rather than the rare symptom.[3] Through the associations between symptoms or the presence of the context in the test phase a preference for the more frequent disease given the ambiguous symptom could also be predicted. These extensions would again handle Experiment 1 but would then fail in Experiments 2 and 3, unless it is argued that the learning mechanism changes based on complexity. However, this seems to be an unusual theoretical move for a theory that generally tends to postulate a fixed basic learning mechanism which should not be affected by greater complexity (see Cobos et al., 2002).

In contrast, for Experiment 1 causal-model theory gives a straightforward account which postulates that people are capable of acquiring adequate causal model representations that contain information about causal structures and their parameters (including the base rate parameter)(see Gopnik et al., 2004; Glymour, 2001; Hagmayer, Sloman, Lagnado, & Waldmann, in press; Lagnado et al., in press; Waldmann & Hagmayer, 2005). Experiment 1

provides the first demonstration that people have the competence to correctly learn about causal parameters irrespective of the experienced sequence of learning information when identical causal models are presented. Causal-model theory also anticipates storage of frequency information, which was found in all experiments.

Although the results of Experiment 1 favor causal-model theory, Experiments 2 and 3 seem to be better predicted by standard associative theories. Basic associative theories predict base rate sensitivity after diagnostic learning, but not after predictive learning. However, strictly speaking the predictions of associative theories for our predictive-learning tasks require an extension. So far, associative theories have not been applied to learning tasks in which learning sequence and test sequence do not coincide. Under the assumption that learners transfer the associative weights from predictive learning to diagnostic testing these theories predict the asymmetry of base rate use, however. One aspect that presents difficulties for this class of theories is the fact that people still encoded base rate information even in the more complex tasks. A possible route for associative theories would be to postulate multiple systems, a frequency based learning system that is used for complex inferences (e.g., against the learning direction), and an associative system for basic tasks that require inferences in the cue-outcome direction (see Price & Yates, 1995, for such a proposal). However, this model does not predict the differences between Experiment 1 and Experiments 2 and 3. An alternative would be to postulate associative learning for complex tasks and restrict causal-model learning to simpler domains (Cobos et al., 2002; Tangen & Allan, 2004). Although this theory is a theoretical possibility, a precise model that incorporates both learning components along with assumptions about the conditions that trigger the chosen learning strategy has yet to be developed (see also López, Cobos, & Caño, 2005).

Causal-model theory also needs to be extended to account for the results of Experiments 2 and 3. Previous research with other tasks has already shown that the competence to form adequate causal representations may break down when the task surpasses the processing

limits of learners (De Houwer & Beckers, 2003; Reips, 1998; Waldmann & Walker, 2005). A plausible hypothesis accounting for performance deficits might postulate that learners confronted with complex tasks give up the goal to construct complete causal model representations that can be flexibly accessed. A complete causal model representation contains knowledge of causal structures and knowledge of the size of the parameters. Since structure knowledge is arguably more important than parameter knowledge, a plausible hypothesis is that learners are mainly interested in learning the structure of the causal models and only focus on the parameters they need for error free performance. For other parameters they might fill in default values or use default estimation strategies (Waldmann & Walker, 2005). The results of Experiments 2 and 3 seem consistent with the notion that participants tend to choose a representation during learning that reduces errors (see also Lovett & Schunn, 1999). Whereas the diagnostic learning task can only be mastered when base rates are taken into account at least implicitly, the predictive learning context permits error free performance without having to use base rate information. Thus, in the predictive learning condition a plausible prediction is that people default on the assumption of equal base rates.

In some respects this proposal is similar to the idea of postulating two learning mechanisms, a causal-model learning mechanism and an associative mechanism (Cobos et al., 2002; López, Cobos, & Caño, 2005; Tangen & Allan, 2004). The basic difference is that the extended causal-model theory does not postulate two separate systems with different learning strategies but a unified learning mechanism that predicts that learners attempt to form causal model representations. Instead of postulating a second mechanism, the main hypothesis is that learners in complex tasks may neglect individual parameters of the causal models that seem less relevant for the present task. Such a model is more parsimonious than a multiple system account, and also has the advantage of postulating a unified probability learning mechanism instead of having to switch between associative and probability learning (see also Gopnik et

al., 2004; Waldmann & Martignon, 1998). However, at this point the data do not permit us to empirically decide between the competing accounts.

The finding that people encode base rates but do not always use them also poses a challenge for causal-model theory. A possible explanation is that people do not combine separately stored base rate information with causal strength information (as required by the Bayes inversion formula) but rely on direct estimates of conditional probabilities. More specifically, in the diagnostic task learners would primarily pick up probabilities (or frequencies) conditional on symptoms, whereas in the predictive task, probabilities would be learned conditional on diseases. Thus, as in the case of associative theories, base rates would be implicitly learned in the diagnostic task, as they are embodied in the probabilities of the diseases conditional upon the symptoms. Because base rates are implicitly embodied in the diagnostic conditional probabilities (see also Gigerenzer & Hoffrage, 1995), the explicit storage of frequencies would be a side effect of probability learning. This would explain why frequencies can be stored but still be neglected in the conditional probability estimates. This finding also places constraints on causal-model theory's account of the good performance in Experiment 1. Apparently learners need to be able to estimate conditional probabilities in both directions to arrive at adequate judgments. Simple knowledge of the base rates is probably not sufficient for good performance.

An interesting question for future research will be whether the competency to acquire flexibly accessible knowledge (as evidenced in Experiment 1) is based on the learning phase or on the retrieval phase. One possibility is that learners attempt to simultaneously learn knowledge in the predictive and diagnostic direction regardless of the learning task when the complexity of the task permits it. Another possibility is that the learners are capable of storing patterns of frequencies and co-occurrences in simple situations as in the present experiment with a single deterministic M-structure, and derive the necessary conditional probabilities from this knowledge base in the test phase (i.e., the retrieval stage).

In summary, none of the competing theories is currently developed far enough to simultaneously account for the symmetries and asymmetries of base rate use in the three experiments. Hopefully future research will show which of the outlined theoretical possibilities is adequate.

Practical consequences

Our research is of considerable practical significance for educational settings. It shows that the philosophy of many medical text books to present information organized around causes (see Thagard, 2000) may lead to deficits when this knowledge has to be used. Base rate neglect with verbally described materials has amply been documented. However, the present findings show that even feedback-based trial-by-trial learning and direct observations of frequency information are not immune to this error. Regardless of how our empirical findings will be theoretically explained, they provide important constraints for the selection of suitable learning and training contexts in education.

References

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35*, 303-313.

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

Christensen-Szalanski, J. J. J., & Beach, L. R. (1982). Experience and the base-rate fallacy. *Organizational Behavior and Human Performance, 29*, 270-278.

Cobos, P.L., López, F.J., Cano, A., Almaraz, J., & Shanks, D.R. (2002). Mechanisms of predictive and diagnostic causal induction. *Journal of Experimental Psychology: Animal Behavior Processes, 28*, 331-346.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theory on human contingency learning. *Quarterly Journal of Experimental Psychology, 55B*, 137-151.

De Houwer, J., & Beckers, T. (2003). Secondary task difficulty modulates forward blocking in human contingency learning. *Quarterly Journal of Experimental Psychology, 56B*, 345-357.

De Houwer, J., Beckers, T., & Vandorpe, S. (2005). An inferential reasoning account of cue competition in human contingency learning. *Learning & Behavior, 33, 2*, 239-249.

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* (pp. 249-267). New York: Cambridge University Press.

Fiedler, K., Brinkmann, B., Betsch, T., & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond base rate neglect and statistical format. *Journal of Experimental Psychology: General, 129,* 399-418.

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

Gigerenzer, G, Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception & Performance, 14*, 513-525.

Gluck, M. A., & Bower, G. H. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, *27*, 166-195.

Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA: MIT Press.

Goodie, A. S., & Fantino, E. (1995). An experientially derived base-rate error in humans. *Psychological Science, 6*, 101-106.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111*, 3-32.

Hagmayer, Y., Sloman, S. A., Lagnado, D. A., & Waldmann, M. R. (in press). Causal reasoning through intervention. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation*. Oxford: Oxford University Press.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*, 356-388.

Holyoak, K. J., & Spellman, B. A. (1993). Thinking. *Annual Review of Psychology, 44*, 265-315.

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

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*, 1-53.

Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 3-26.

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (in press). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation*. Oxford: Oxford University Press.

López, F. J., Cobos, P. L., & Caño, A. (2005). Associative and causal reasoning accounts of causal induction: Symmetries and asymmetries in predictive and diagnostic inferences. *Memory & Cognition, 33*, 1388-1398.

Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability, and base-rate neglect. *Journal of Experimental Psychology: General, 128,* 107-130.

Medin, D. L., & Edelson, S. E. (1988). Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General, 117*, 68-85.

Price, P. C., & Yates, J. F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1639-1655.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: an essay on the cognitive unconscious.* New York, NY: Oxford Univ. Press.

Reips, U.-D. (1998). *Erwerb von Kausalwissen: Lernrichtung und Diagnostizität von Effekten* [Acquisition of causal knowledge: learning direction and diagnosticity of effects]. St. Augustin: Gardez!

Shanks, D. R. (1990). Connectionism and the learning of probabilistic concepts. *Quarterly Journal of Experimental Psychology, 42A*, 209-237.

Shanks, D. R., & López, F. J. (1996). Causal order does not affect cue selection in human associative learning. *Memory & Cognition, 24*, 511-522.

Spalding, T. L., & Murphy, G. L. (1999). What is learned in knowledge-related categories: Evidence from typicality and feature frequency judgments. *Memory & Cognition, 27*, 856-867.

Tangen, J. M., & Allan, L. G. (2004). Cue-interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition, 32*, 107-124.

Thagard, P. (2000). *How scientists explain disease*. Princeton: Princeton University Press.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (pp. 49-72). Hillsdale: Erlbaum.

Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 47-88). San Diego: Academic Press.

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 53-76.

Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychological Bulletin & Review, 8*, 600-608.

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 216-227.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, *121*, 222-236.

Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General*, *124*, 181-206.

Waldmann, M. R., & Martignon, L. (1998). A Bayesian network of causal learning. In M. A. Gernsbacher & S. J. Derry, *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 1102-1107). Mahwah, NJ: Erlbaum.

Waldmann, M. R., & Walker, J. M. (2005). Competence and performance in causal learning. *Learning & Behavior, 33,* 211-229.

Yamauchi, T., Love, B. C., & Markman, A. B. (2002). Learning non-linearly separable categories by inference and classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 585-593.

Yamauchi, T., & Markman, A. B. (1998). Category-learning by inference and classification. *Journal of Memory and Language, 39,* 124-148.

Author Note

All experiments were planned and conducted when both authors were affiliated with the University of Tübingen, Germany. We would like to thank our helpers in conducting the experiments, R. Decker, T. Härlen, Y. Hagmayer, B. Hanfstein, H. Kloos, M. Schwemmle, and M. Buehner, M. Vadillo and an anonymous reviewer for helpful comments on an earlier version of this manuscript. Correspondence concerning this article should be addressed to U.-D. Reips, Psychologisches Institut, Universität Zürich, Binzmühlestrasse 14/13, 8050 Zürich, Switzerland, e-mail: u.reips@psychologie.unizh.ch, or to M. Waldmann, Institut für Psychologie, Universität Göttingen, Gosslerstr. 14, D-37073 Göttingen, Germany, e-mail: michael.waldmann@bio.uni-goettingen.de.

Footnotes

[1]After the diagnostic ratings, which always were requested first, participants in some of the experiments were asked for their assessment of the predictive relationships (i.e., the probability of symptoms given the individual diseases). Since these data are of little theoretical significance in the present context we will not discuss these results here.

[2]We have also conducted an experiment with two M-structures (4 diseases, 6 symptoms), which will not be reported here because the results are very similar to the results of the present Experiments 2 and 3.

[3]These theoretical possibilities were suggested by M. Buehner and M. Vadillo.

List of Figure Captions

Figure 1. The same causal structure can either be learned from causes to effects or from effects to causes.

Figure 2. Predictive versus diagnostic learning of a single M-structure with a frequent and a rare cause. The disease pleroia is three times as frequent as the disease spetitis, shortness of breath (ambiguous symptom) is caused by both diseases. Unique symptoms are caused by only one of the diseases.

Figure 3. Schematic description of the learning task. The final test requests access to the diagnostic direction.

Figure 4. Mean probability ratings for the frequent and rare diseases in the presence of the ambiguous symptom in Experiment 1 (single M-structure). Data labels show differences of ratings between frequent and rare diseases given the ambiguous symptom.

Figure 5. Mean estimates of frequencies of frequent and rare diseases after diagnostic and predictive learning in Experiment 1 (single M-structure). Data labels show differences of estimates between frequent and rare diseases given the ambiguous symptom.

Figure 6. Mean probability ratings for the frequent and rare diseases in the presence of the ambiguous symptoms in Experiment 2 (triple M-structure).

Figure 7. Mean estimates of frequencies of frequent and rare diseases after diagnostic and predictive learning in Experiment 2 (triple M-structure).

Figure 8. Percent of diagnoses of frequent and rare diseases causing the ambiguous symptoms (i.e., implicit test of base rate use) in Experiment 3 (triple M-structure). Data labels show differences of percentages.

Figure 9. Mean estimates of frequencies of frequent and rare diseases after diagnostic and predictive learning in Experiment 3 (triple M-structure).

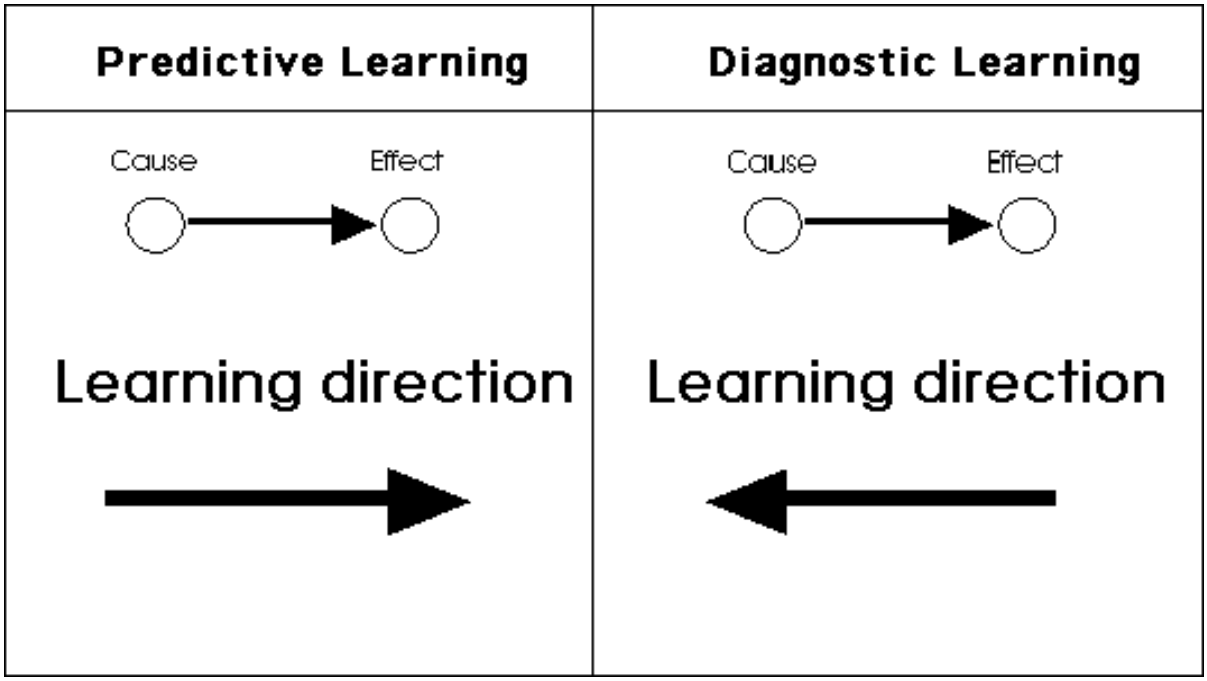

Figure 1.

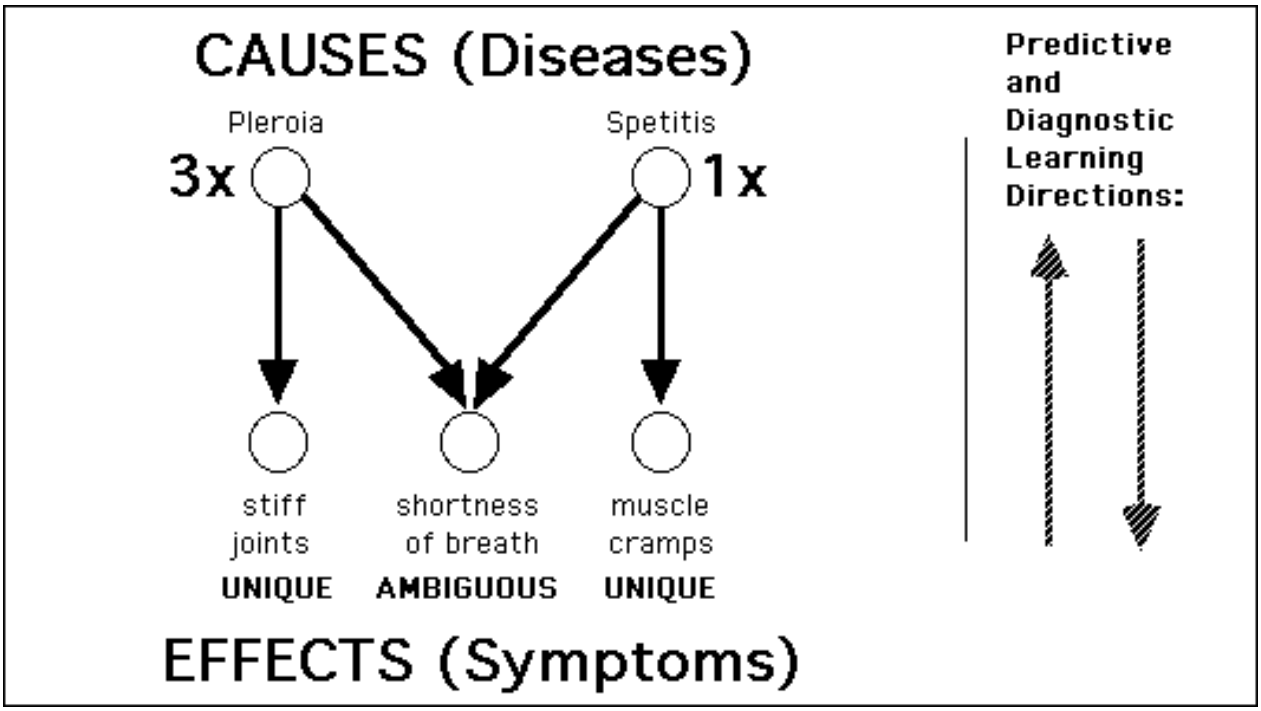

Figure 2.

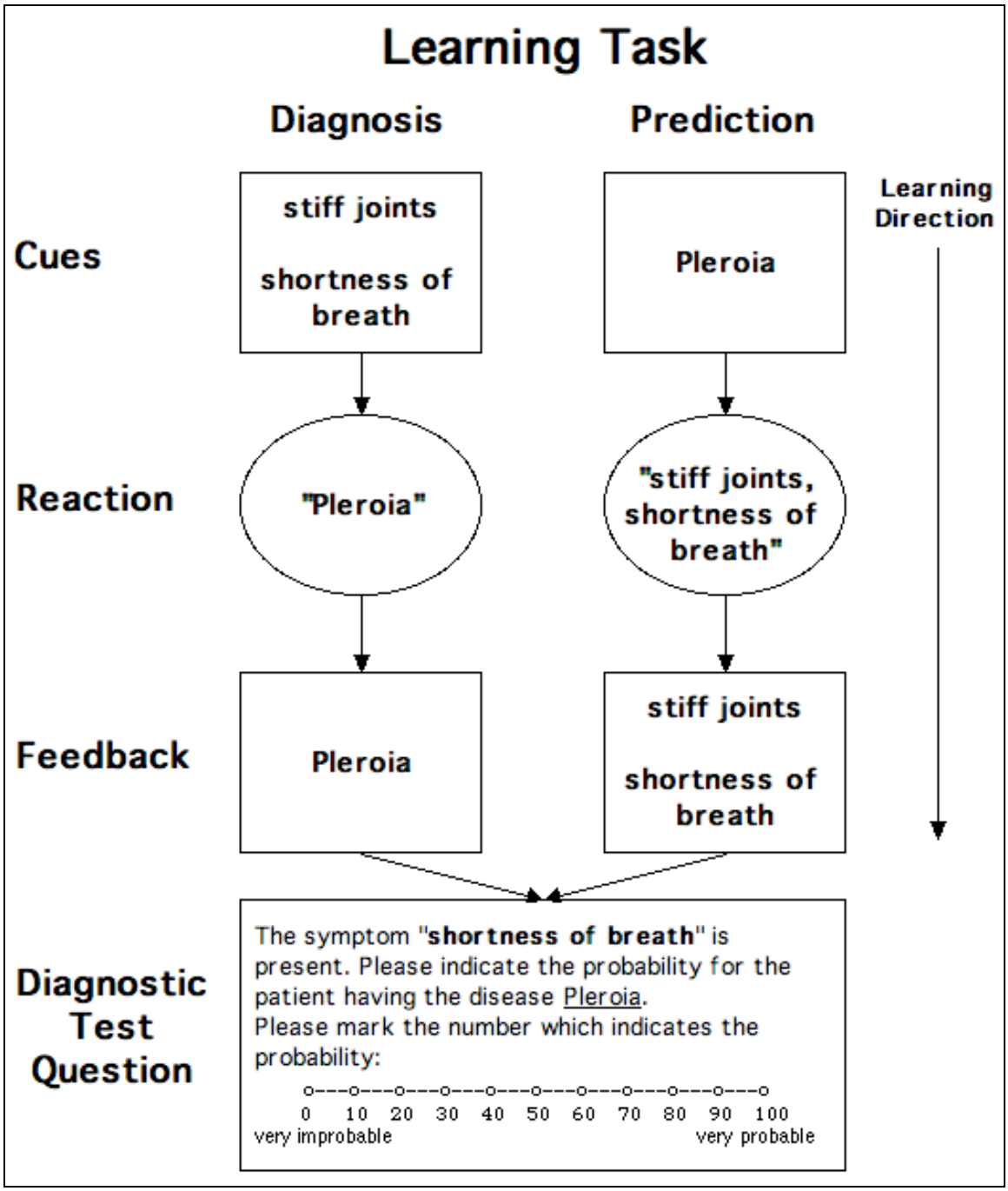
Learning Task
Diagnosis
Prediction
Cues
stiff joints
shortness of breath
Pleroia
Learning Direction
Reaction
"Pleroia"
"stiff joints, shortness of breath"
Feedback
Pleroia
stiff joints
shortness of breath
Diagnostic Test Question
The symptom "shortness of breath" is present. Please indicate the probability for the patient having the disease Pleroia.
Please mark the number which indicates the probability:
o---o---o---o---o---o---o---o---o---o---o
0 10 20 30 40 50 60 70 80 90 100
very improbable
very probable

Figure 3.

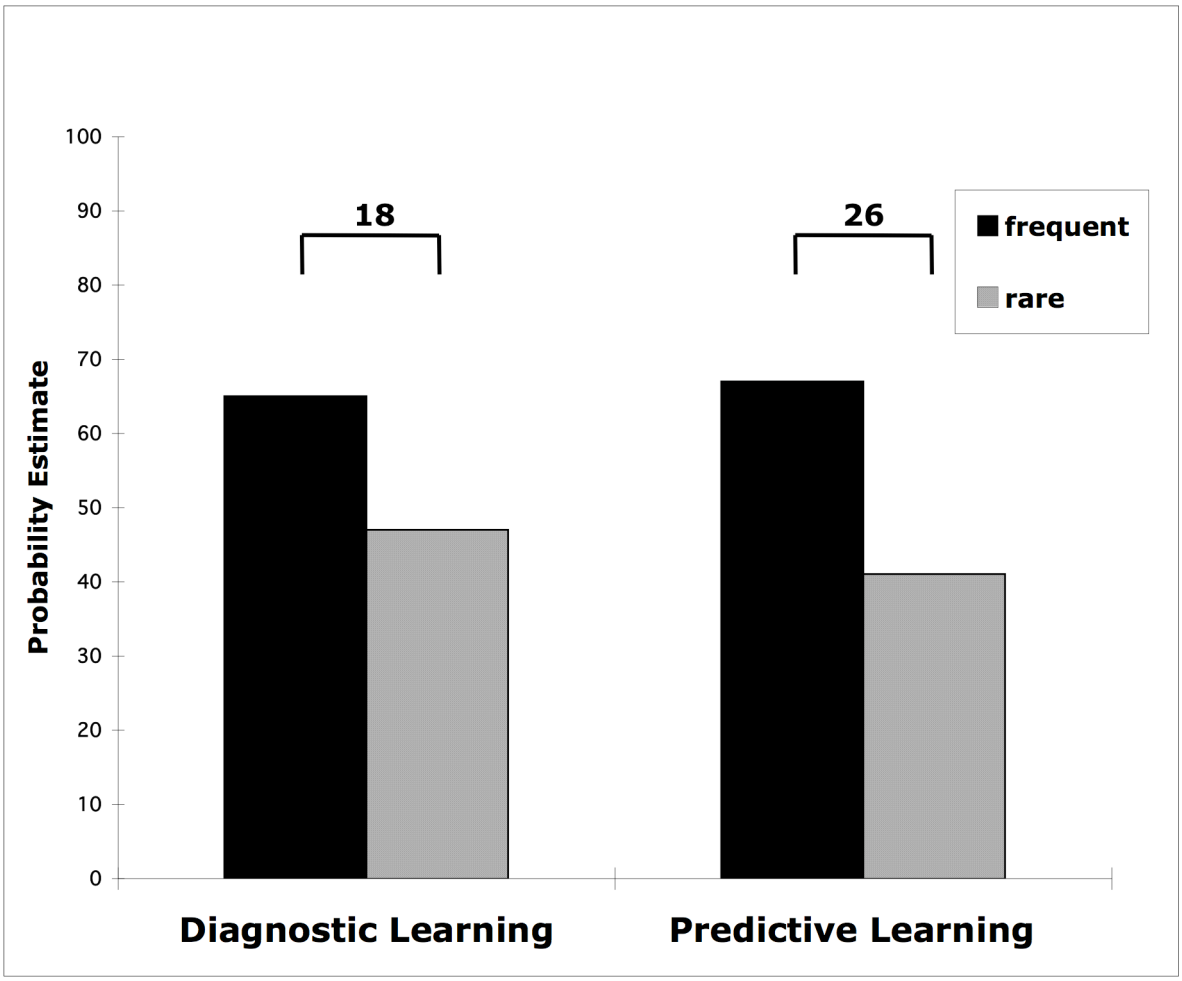
100
90
80
70
60
50
40
30
20
10
0
Probability Estimate
18
26
frequent
rare
Diagnostic Learning
Predictive Learning

Figure 4.

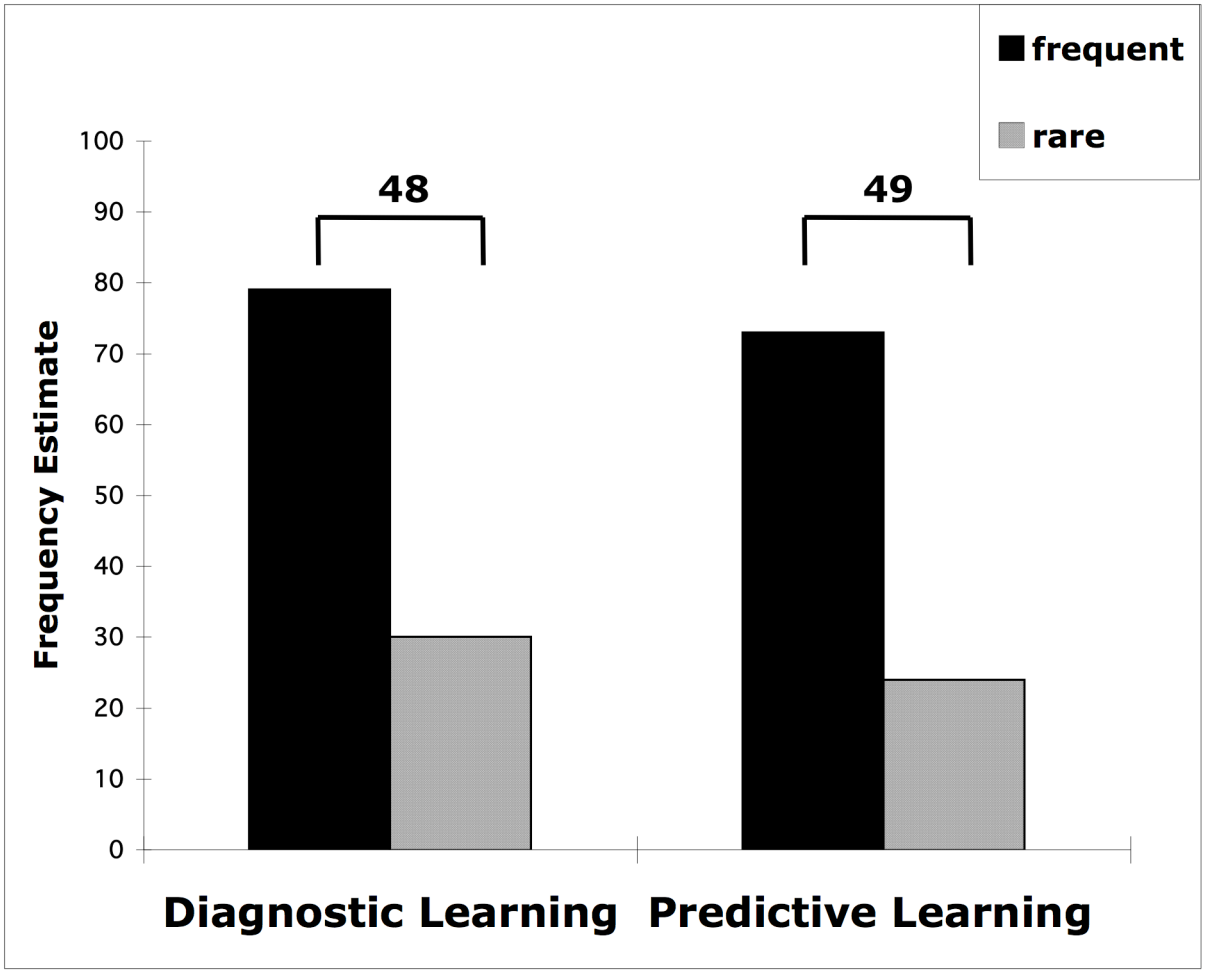
frequent
rare
100
90
80
70
60
50
40
30
20
10
0
Frequency Estimate
48
49
Diagnostic Learning
Predictive Learning

Figure 5.

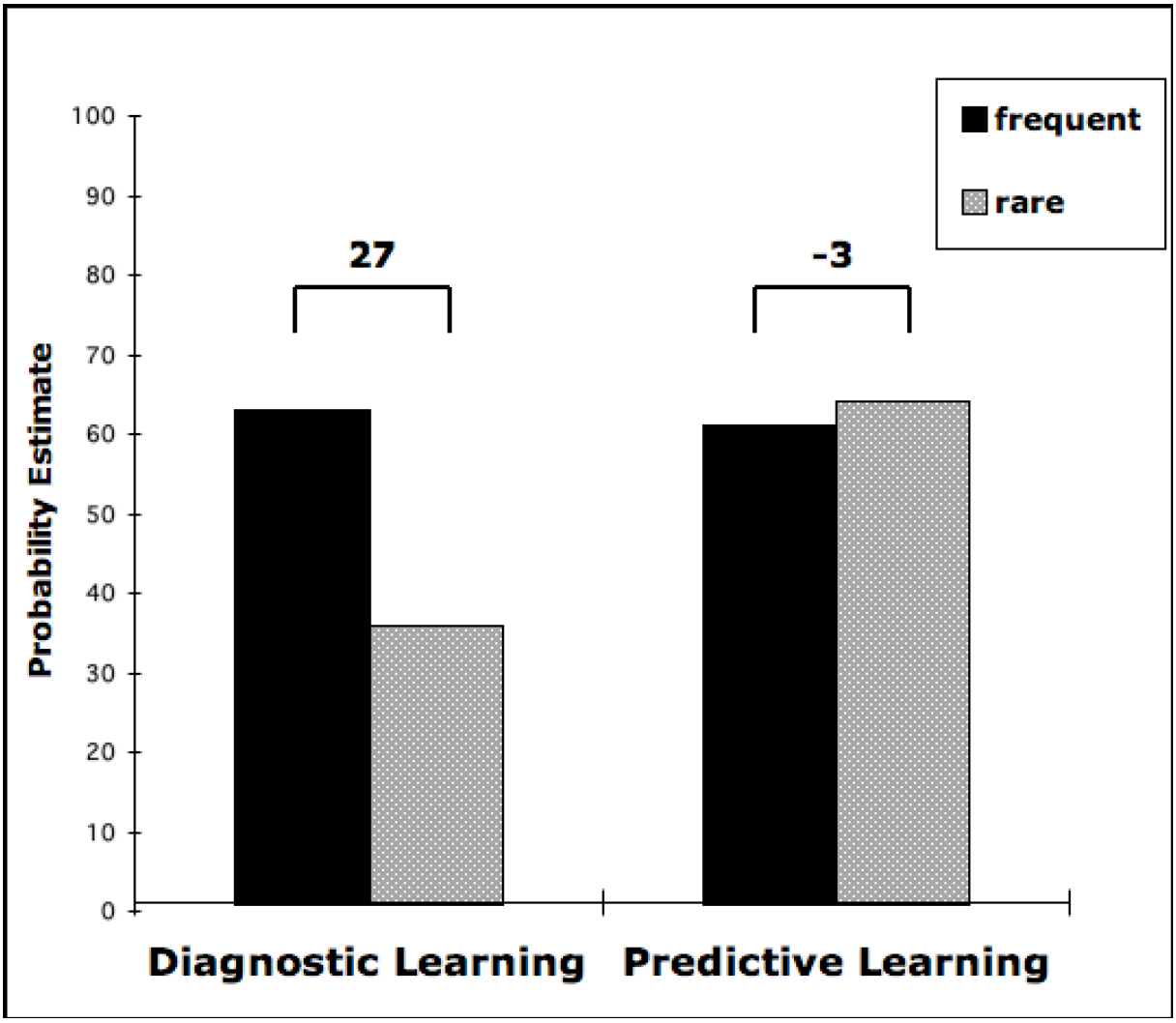
frequent
rare
27
-3
Probability Estimate
100
90
80
70
60
50
40
30
20
10
0
Diagnostic Learning
Predictive Learning

Figure 6.

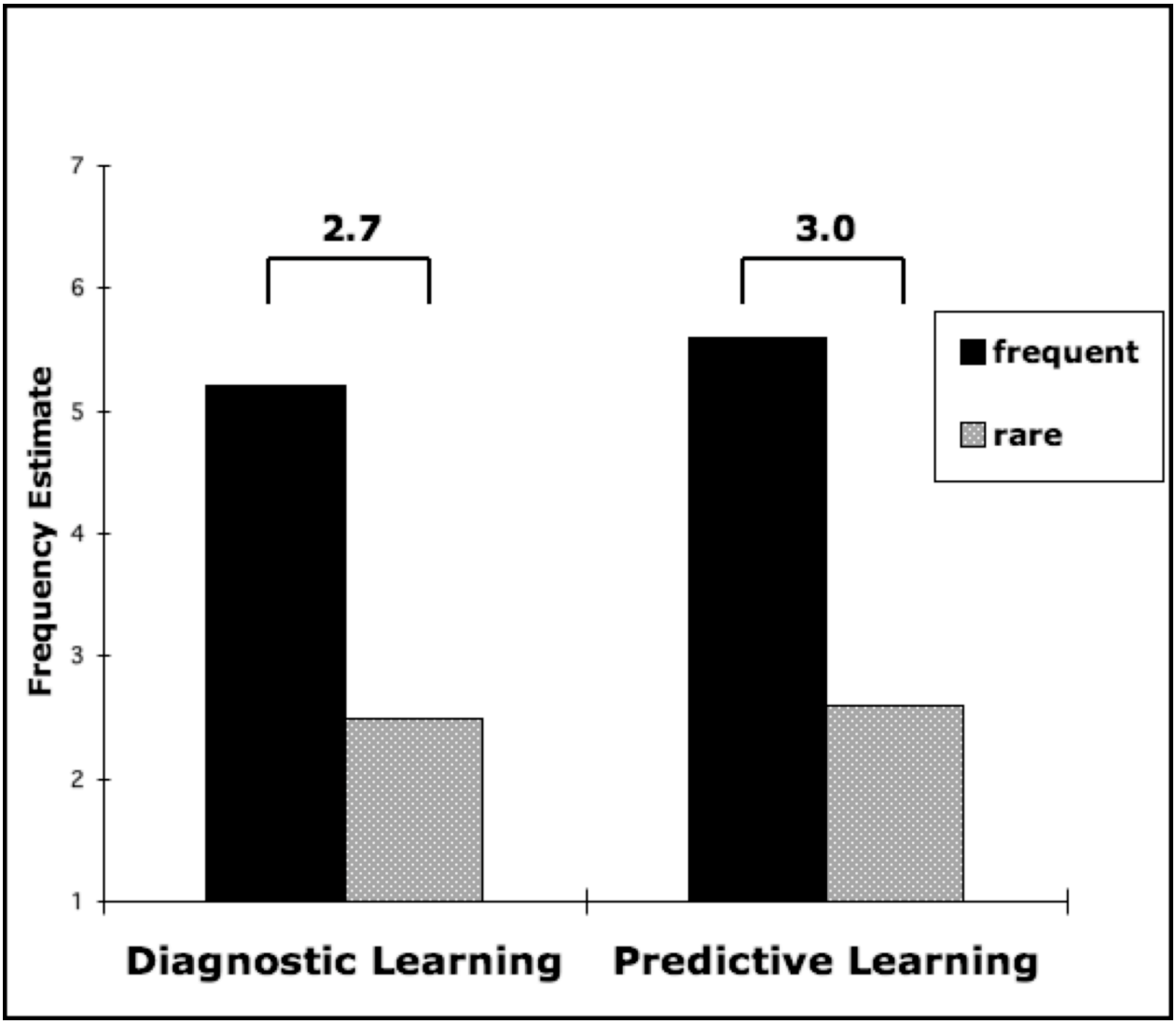
7
6
5
4
3
2
1
Frequency Estimate
2.7
3.0
frequent
rare
Diagnostic Learning
Predictive Learning

Figure 7.

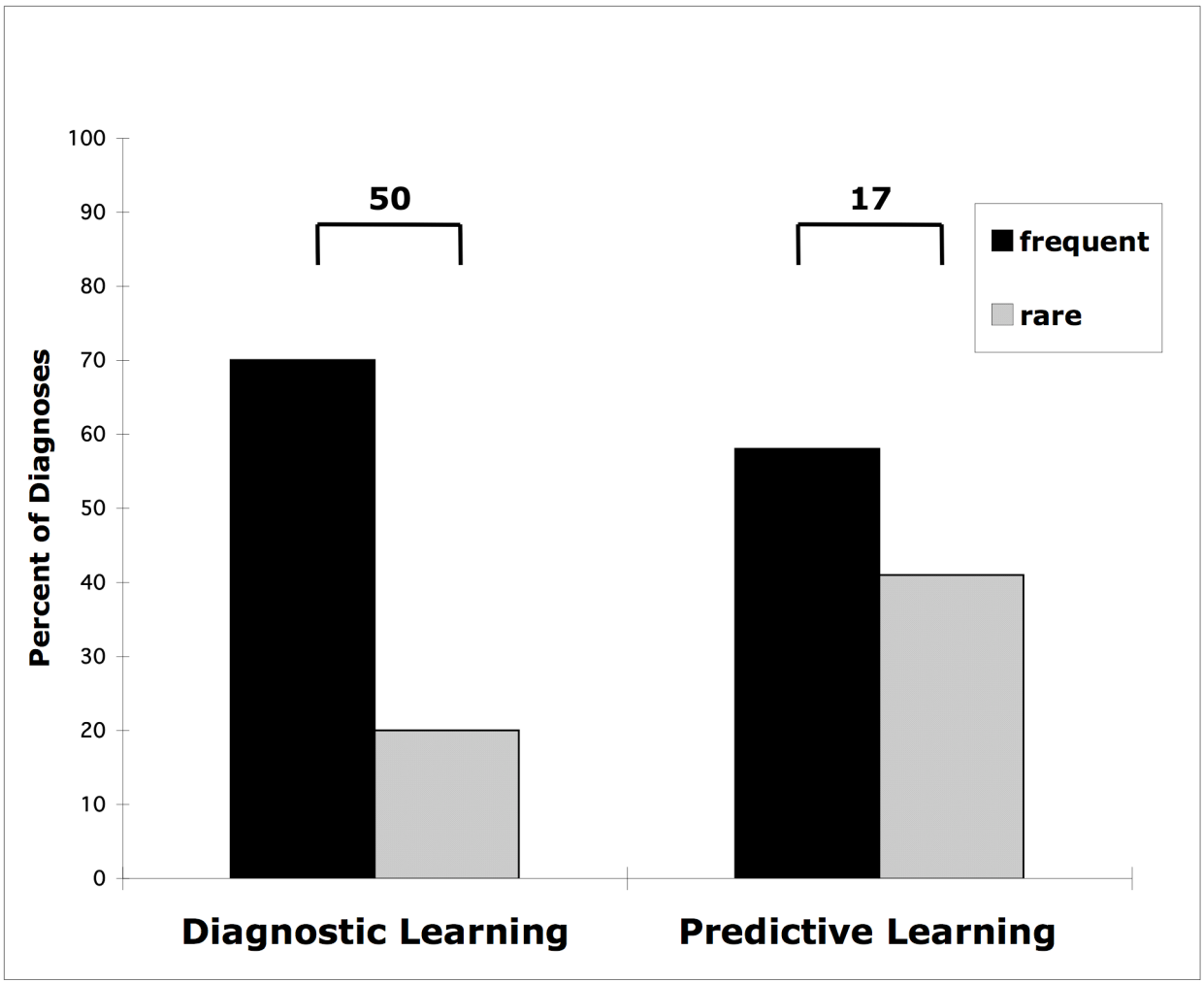
100
90
80
70
60
50
40
30
20
10
0
Percent of Diagnoses
50
17
frequent
rare
Diagnostic Learning
Predictive Learning

Figure 8.

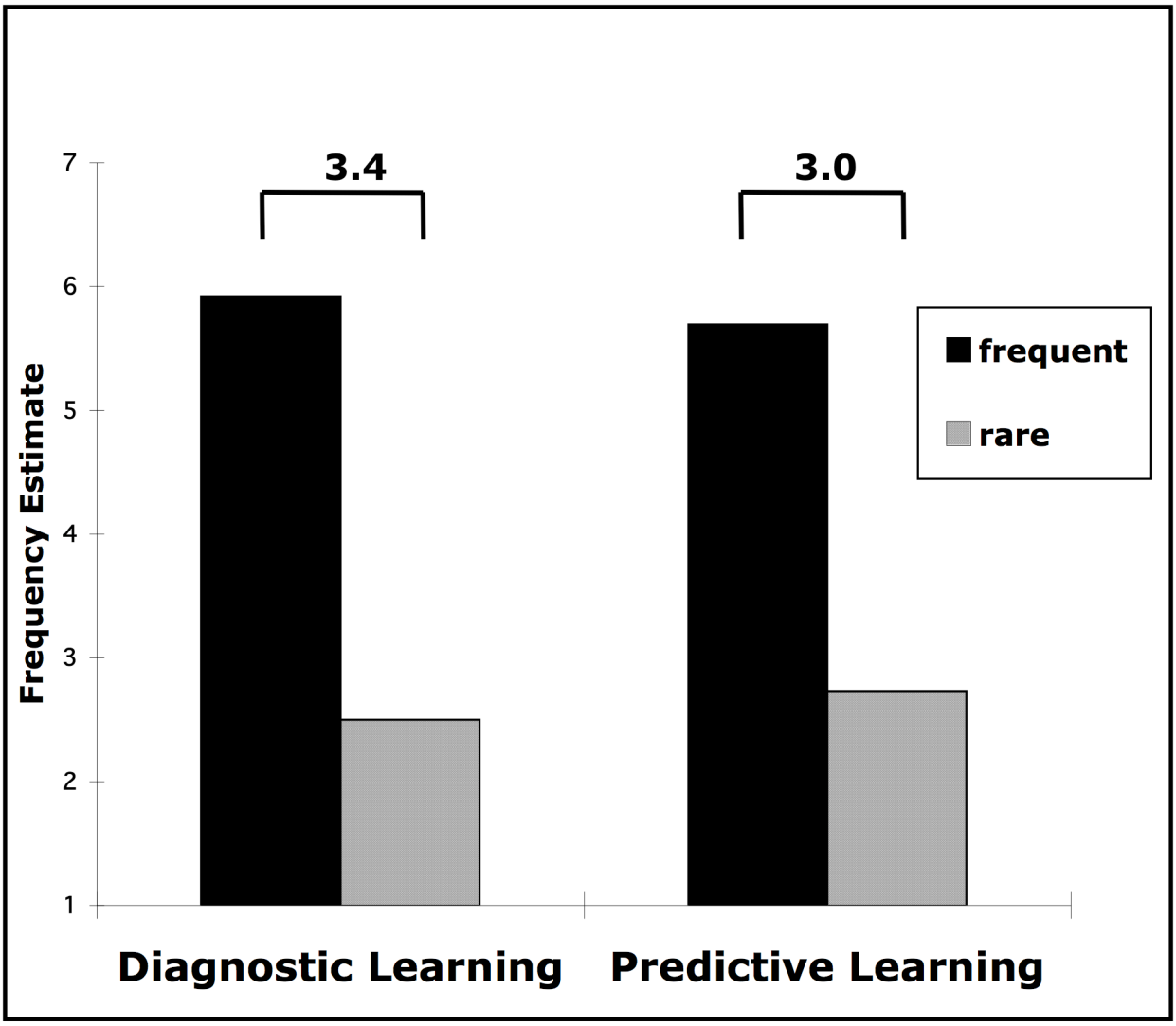
3.4
3.0
7
6
5
4
3
2
1
Frequency Estimate
frequent
rare
Diagnostic Learning
Predictive Learning

Figure 9.