Logical Patterns in Individual and General Predication

Momme von Sydow (momme.von-sydow@psychologie.uni-heidelberg.de
University of Heidelberg, Department of Psychology, Hauptstr. 47-51,
D-69117 Heidelberg, Germany

Abstract

Probability judgments about logical propositions have raised substantial doubts about human rationality. Here we explore the idea that people's probability judgments often may not refer to the relative frequency of a set, but instead to the probability of an explanatory logical pattern given the data. This idea has been formalized by Bayesian logic (BL), predicting a system of frequency-based logical inclusion fallacies. The studies presented concentrate on comparing probability judgments about sentences logically relating two attributes of a class or an individual (humans, animals, artifacts). Although BL cannot model probabilities of individual predications directly, it can do so if one assumes that inferences are made about unknown individuals based on imagined samples. The results for general as well as individual predication show a high number of systematic inclusion fallacies in line with BL. Nevertheless, some deviations were found. In the General Discussion, a polycausal approach to inclusion fallacies is advocated. In addition, even if pattern probabilities seem to play a major role, it is suggested that extensions of the BL model may be needed to account for further aspects of real-life predication. Overall, however, even the basic BL model was surprisingly successful for predicting probability judgments about general as well as individual predications.

Keywords: Probability judgment; bias; conjunction fallacy; inclusion fallacy; inductive logics; predication.

Narrow Norms of Predication?

Throughout Western philosophy (Aristotle, the Stoics, Leibniz, cf. even Kant and Hegel), and particularly since logical positivism (Frege, Wittgenstein, Russell, Whitehead), logic has been central (with slightly different understandings) to defining standards of rational thought. Today, standard calculi of logic and probability may appear narrow in comparison to the much broader Greek concept of *logos*, but they provide a rigidly defined standard of rational thought. And yet there is much evidence that people's actual reasoning seems to violate these basic calculi. Thus psychology is torn between the Scylla of abandoning normative reasoning (e.g., psychologism) and the Charybdis of claiming that people are fundamentally irrational, even with regard to the simplest rules of these calculi. Although there seems to be some truth in Kahneman and Tversky's (1996) warning against "normative agnosticism", the arguments of Gigerenzer and colleagues (e.g., Gigerenzer, 1996) against them seem reasonable as well: that is, that the blind application of the "narrow norms"' of logic and probability-theory often seem misguided. In my view, a domain-specific understanding of rationality may allow for a middle course between these positions. Context-sensitive norms of reasoning that account for our goals as well as the precondition of our models may not need to give up the core of the concept of rationality (cf. von Sydow, 2011).

When the calculi of logic and probability are applied in psychology, standard logic is normally used in deductive, and standard probability theory in inductive contexts. Here we consider both in assessing the inductive probability of logical relationships. Propositional logic addresses the combination of atomic propositions (that can either be true or false) with connectives (AND, OR, EITHER OR, NEITHER NOR, etc.). In the tasks we investigate probability judgments involving several alternative logical sentences, with different logical connectives relating two properties. We are either concerned with the properties of an entity or of a class of entities (individual vs. general predication).

The suggested domain-specific approach to rationality should consider the context and the goals implied. The context of our probability-judgment task is the assignment of attributes to a class. What is a reasonable, observation-based norm for predicating specific logical relationships between attributes, and how does this relate to probabilities (von Sydow, 2011)? At first sight, propositional logic seems a plausible candidate. A sentence such as "ravens are black and they can fly" logically seems to predicate the conjunction of attributes ($B \wedge F$) to the class of ravens (R). From a falsificationist perspective, this predication is valid as long as no single exception defies the rule. Predications about contingencies in the actual world (in contrast to mathematics) would all be rendered false, since one may assume that they are not free of exceptions. For instance, albino ravens exist, as well as other non-black ravens. It therefore seems reasonable to replace a purely logical adequate criterion of predication by a high-probability criterion (cf. Schurz, 2005). In the raven example, correct predication would require that $P(B \wedge F|R) > \psi$, with ψ being the high-probability criterion. This proposal additionally appears to solve the problem of non-monotonicity, since now an adequate predication may become inadequate (and vice-versa) during further data-sampling. Nonetheless further problems remain.

Here only the problem of set-inclusion is sketched (cf. von Sydow, 2011, von Sydow & Fiedler, 2012). The frequentist/extensional probability of the predication "ravens are black and they can fly" can never be larger than the probability of the inclusive disjunction P("ravens are black or they can fly or both") ($P(B \vee F|R)$), since the former refers to a subset of the latter. Likewise, the AND sentence cannot have a larger probability than the tautology (P(all feature-combinations are possible)). Using an extensional probability-criterion excludes preferring the predication of a more specific hypothesis over a (less informative) more general one. The tautology ($P(B \text{ T } F|R) = 1$) would always be a rational predication, even independent of data. Therefore, extensional probabilities could not be reasonable evidence-based criteria for adequate predication.

Probabilities of Noisy-Logical Patterns

One way to resolve this problem and the problem exceptions together is to assume that people tend to judge the probability of alternative explanatory logical patterns instead of the relative size of particular sets, when concerned with probabilities of alternative logical predications, each meant as an explanation of the whole situation. A first formalization of this idea has been provided by von Sydow (2011, cf. von Sydow & Fiedler, 2012). Here only the idea of the model, called Bayesian Pattern Logic (Bayesian Logic, or BL), is sketched, without providing a formal model. In the wake of the renaissance of Bayesian models (cf. Chater, Tenenbaum, Yuille, 2006; Kruschke, 2008; Oaksford & Chater, 2007) it is formulated as a Bayesian approach. It formalizes the idea of explanatory logical patterns (an AND-pattern, an EITHER-OR-pattern, etc.), under absence of further factors. The model provides the measure of fit between a 2×2 frequency table input and 2×2 probability tables that may hypothetically have produced the data (hypothetical noisy-logical explanations). The probability tables are based on logical truth tables assuming equiprobability of true cases (cf. Johnson-Laird et al., 1999; Tenenbaum & Griffith, 2001) and a uniform noise function. Based on these basic assumptions, the model first establishes the likelihood that some observed data have been produced by the probability tables, $P(D|PT)$. To obtain the posteriors, the probabilities of these hypothetical noisy-logical explanations given the data ($P(PT|D)$), one uses the Bayes theorem. To obtain the probability of a connective, one sums up the corresponding posteriors over all noise levels (for technical details, see von Sydow, 2011; cf. von Sydow, 2009, von Sydow & Fiedler, 2012).

In sum, the extensional probability of a set (relative frequency) is here replaced by the second-order probability of noisy-logical patterns of probabilities (all four cells of a PT add up to 1). These patterns serve as hypothetical logical explanations. It is predicted that people use pattern probabilities to explain a whole situation in logical terms (class *X* is *A* and *B*), instead of judging the size of a set or subset. Accordingly, *P*(ravens are black and they can fly) should be high, not because there are few exceptions but because our subjective frequency pattern best fits a noisy AND-pattern. If one is concerned with pattern probabilities, the probability that a data-set may be produced by an AND-pattern may well be higher than that for an OR-pattern: $P_P(B \wedge F|R) > P_P(B \text{v} F|R)$. By contrast, a narrow application of extensional probability always requires that $P_E(B \wedge F|R) \leq P_E(B \text{v} F|R)$ (cf. von Sydow, 2011).

Previous work in the conjunction-fallacy debate generally concerned a quite different, story-based task, showing that people may judge the conjunction more probable than the conjunct, e.g., $P(B \,\&\, F) > P(B)$ (Tversky & Kahneman, 1983). In a few cases, CFs were also shown without stories (e.g., Lagnado & Shanks, 2002). In any case, most authors have assumed that such conjunction-judgments involve a "conjunction fallacy" (CF). Conversely, BL suggests a rational explanation at least of a particular class of CFs [for convenience they are nonetheless called "fallacies" here].

The application of BL led to several new predictions and corroborative findings—for instance, on double CFs, sample-size effects, and pattern-sensitivity effects (von Sydow, 2011). The concept of CFs has been generalized to apply to system of logical connectives based on summary information (von Sydow, 2009) or sequential input (von Sydow, 2012). Whether or not other theories may account for independent causes of CFs (e.g., Lagnado & Shanks, 2002; Tentori, Crupi, Russo, 2012), these results could not be explained by any other current theory. It seems plausible, then to conclude the existence of a class of pattern-based CFs. Additional factors—for instance, unclear set-inclusions (Sloman, Over, Slovak, & Stibel, 2003), illicit implicatures (Hilton, 1995; cf. Hertwig et al. 2008), and probability format (Fielder, 1988)—remain plausible further facilitators for CFs, even if one is concerned with extensional probability judgments. Nevertheless, a high proportion of CFs were found even when simultaneously using clear formulations, clear set-inclusions, rating scales, and frequency information (von Sydow, 2011).

Individual vs. General Predications Based on Real-Life Frequencies

The investigations reported here address three issues.

(1) Previous tests of BL used explicit frequency inputs, either in a table format (von Sydow, 2011) or in an experienced sequential learning format (von Sydow & Fiedler, 2012). Although this allowed for precise tests of plausible models, it may differ from real-life predication where samples often have to be retrieved from memory. Moreover, the explicit frequencies presented in other tasks might have suggested the use of something like BL. We therefore assess here subjective frequencies of real-life predication independently from the task where participants judged probabilities of different logical sentences. Whereas previous tests focused on the variation of frequencies and only used a small number of scenarios, in order to reduce the influence of uncontrolled priors or other disruptive factors, we here used several different scenarios involving people, animals and artifacts.

(2) Despite previous success in modeling frequency-based prediction, it is an open issue whether the pattern idea is applicable to individual predications as well. BL cannot be applied to individual predication without an auxiliary hypothesis. The formal model has a frequency-based input—the four cells of a contingency matrix, ($f(B \& F)$; $f(B \& \neg F)$; $f(\neg B \& F)$; $f(\neg B \& \neg F)$). Although some frequentists have been skeptical about probability judgments in individual cases, it seems plausible that humans often base their probability estimates, even for individual cases, on imagined subjective frequencies. The explored auxiliary hypothesis is that for individual predications (concerning e.g., a raven), one may - in the absence of further information - simply imagine a hypothetical sample of ravens. This may be used as input for BL (suggested by von Sydow, 2011).

Overview

In Preliminary Study 1 we first sampled sentences by asking participants for sentences that related two attributes logically. In Preliminary Study 2, participants provided estimates for the frequencies entered in a contingency table relating these attributes. Then in the main study we investigated general vs. individual predication and assessed in 30 scenarios which logical connective relating two attributes was judged to be most probable (an extended CF task with several connectives). We then modelled the predictions of BL based on the subjective frequencies from Preliminary Study 2. Finally, we compared the model predictions with the results for general and individual predication.

Preliminary Study 1

Participants (twelve students from the University of Göttingen) had to fill in the blanks for 6 sentences, each concerning a different logical relationship between two attributes. For each sentence one filled in a class and two attributes, such as "Normally [In der Regel] ___ are either ___ or _____". The order of the six connectives was permutated, and an example was provided: "Normally chairs have four legs AND (at the same time) they have a seat." The predications consistently employed either the verb "to be" or "to have".

After obtaining the results, we narrowed down the number of sentences for Study 2. We excluded arbitrary sentences, obviously deterministic sentences, and sentences that seemed to contain an unwanted or over-complex causal background. We aimed to focus on the relation between two attributes of a class whose co-occurrence could be described in simple logical terms (e.g., ravens are black and can fly). In addition, we supplied four more sentences.

Preliminary Study 2

Method Participants (23 students from the University of Göttingen, 78 % female) provided subjective frequency judgments for the co-occurrence of two attributes in a 2 × 2 contingency table for each of the 50 scenarios investigated. For each randomized scenario, participants assigned a sample of 100 hypothetical cases to the four cells of a table (*a*, *b*, *c*, *d*; see Figure 1).

Results Figure 2 shows four examples for the 50 resulting

Professional Basketball Players

Imagine **100 professional basketball players**.
How frequently do you think the combinations of attributes in the table occur? Sort the 100 cases into the four attribute combinations, giving each a number.

	Quick	Slow
Taller than 1,80 m	*tall & quick* *a*	*tall & slow* *b*
Shorter than 1,80 m	*short & quick* *c*	*short & slow* *d*

How certain are you that your estimated frequency distribution is roughly valid?

0 1 2 3 4 5 6 7 8 9 10
uncertain certain

Figure 1: Assessment of frequency estimates in Preliminary Study 2.

four-cell frequency distributions. The results were later used as input for BL to predict the probability judgments in the main experiment. Based on this study the scenarios were selected so as to have four scenarios for each of the six focused connectives. Two scenarios predicted the main connective with the highest relative frequency of participants (even if the pattern probability was below 50%). The professional-basket-ball-player scenario (Fig. 1a) is an example for an AND-connective. *P*(tall AND also quick) is expected predominantly to be estimated higher than the probabilities of larger sets (despite exceptions). For two further connectives, the second most frequent hypothesis was predicted almost as often as the first. The application of the schema worked quite well, apart from the OR-class, where all scenarios reflected at best the second noise level. Finally, we investigated six 'noise-scenarios', where the predictions of BL become less clear, favoring even more than two connectives (generally to at least three).

Main Experiment

Method The experiment had a 2 (general predication vs. individual predication, between subjects) × 30 (scenarios, within subjects) design. 20 participants judged for each scenario which of 15 logical sentences connecting two target attributes is most probably valid (extended CF task). The 30 scenarios were presented in random order and concerned people, animals and artifacts.

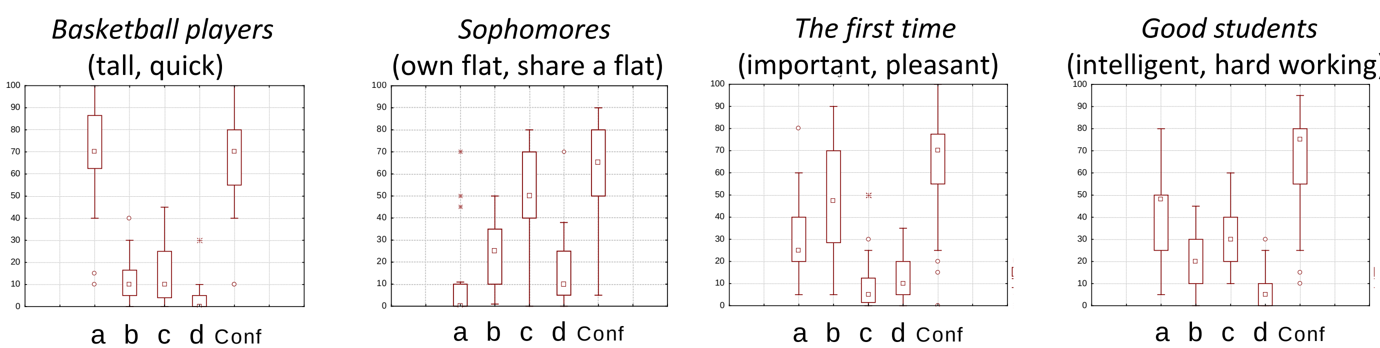

Figure 2: Boxplots depicting the distribution of the estimated frequencies in the four cells of the contingency matric (and of the confidence ratings) for four example scenarios (Preliminary Study 2; Median; 25%-75% boxes)

The instruction for both general and individual predication-conditions followed the same pattern, e.g., “Imagine [hundred / one] professional basketball player[s]. We are concerned with several propositions about [a] professional basketball player[s]. […] Please tick the proposition that in this situation seems most probable to you.” Participants should answer intuitively. In each of the scenarios participants selected the 1 out of 15 that seemed most probable. For the general predication condition, propositions opened with the class (e.g., “Professional basketball players are…”), and in the individual predication condition, with the individual (e.g., “A professional basketball player is…”). The 15 hypotheses always occurred in the same order, referred to all 16 dyadic logical connectives apart from the falsum/contradiction. For instance: *A* AND *B* (H1); *A* AND not-*B* (H2); NEITHER *A* NOR *B* (H4); A (H5); EITHER *A* OR *B* (H9); A OR B OR BOTH (H11), and everything is possible (Verum/Tautology, H15).

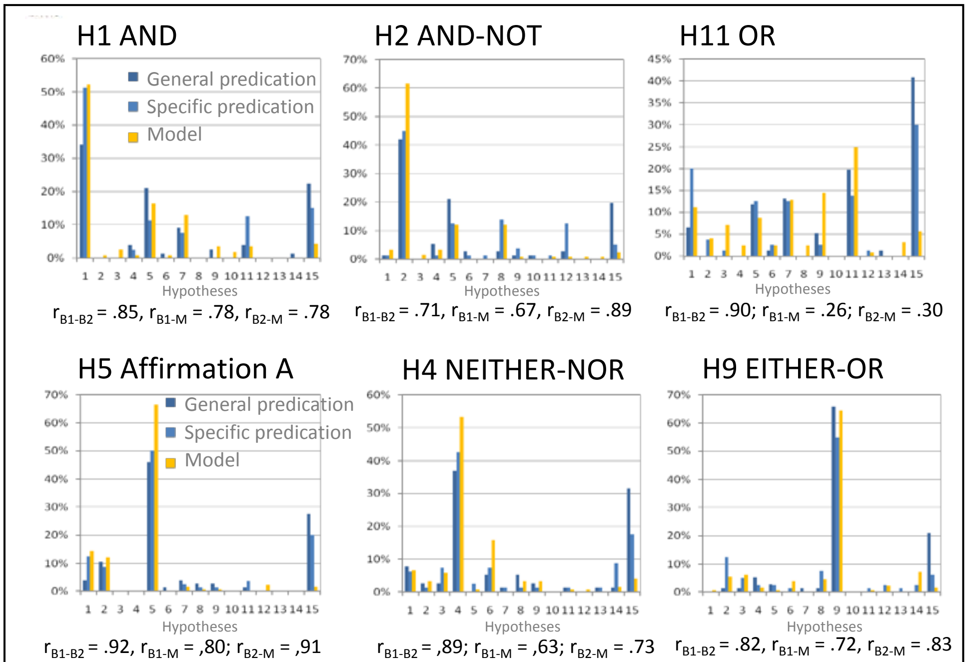

Figure 3: For all six focused types of connectives, the graphs provide a visualization of the results for general and specific predications (percentage of selections) and of the model-predication (averaged over scenarios). On the ordinate, the proportion of actually selected or predicted hypotheses (H1 to H15) is shown (cf. text for details).

CFs ($P(A \wedge B) > P(A)$) may be due to reinterpretation of the logical connectives according to standard conversational implicatures (e.g., Hilton, 1995; Hertwig, Benz & Krauss, 2008). If the affirmation *A* is contrasted with “*A* AND *B*” it may indeed reasonably be represented as “$A\wedge$non-*B*”. To avoid such misunderstandings, we in all studies used the formulation “*X* are *A* (and they are *B* or not-*B*)”. Likewise, “*A* AND *B*” in ordinary language may well refer to “*A* OR *B* (or both).” In this interpretation, $P(A \vee B) > P(A)$ is not fallacious. We used an OR-hypothesis and the following AND-formulation: “*X* are *A* (e.g., taller than 1.8m) and *at the same time B* (e.g., quick).” The verum read: “*X* are tall and quick, tall and slow, short and quick, or short and slow (all combinations).”

Forty participants (from the same population) volunteered to take part, receiving either course-credit or a fee.

Modeling For all scenarios we calculated the predictions of the model based on Study 2. For each participant and scenario we used the estimates for the four cells of the contingency table as input for BL, determining which hypothesis this participant would select as most probable. For reasons of simplicity we ignored further rankings. Calculated for all participants, this provided a reasonable prediction for the distribution of selections in the main task.

Results Figure 3 shows the predictions of the model as well as the accumulated results for the six types of scenario (referring to different dominant connective). Each chart represents four scenarios. Although grouped this way by design, based on Study 2, the scenarios should differ somewhat (low noise vs. high noise). Figure 3 therefore provides merely a simplifying visualization. Nonetheless it does depict the main pattern concisely and well.

In the H1 scenarios (involving, e.g., the basketball-players and ravens scenarios) the AND-connective was the most frequently selected. Such judgments involve estimating *P*(*A* AND *B*) to be more probable than *P*(*A*), *P*(Either both attributes or none), *P*(*A* OR *B* or both), and *P*(Tautology). Given the presence of exceptions (e.g., Figure 2a), this would traditionally be interpreted to involve several simultaneous logical inclusion fallacies (von Sydow, 2009). Moreover, it appears that on this level the overall distribution of selections reflect the predictions quite closely. The most striking deviation in this and the other scenarios, however, was that H15 (the tautology), the extensionally correct solution, was selected more frequently than predicted (cf. Fig. 2).

In the H2 scenarios, participants predominantly selected the predicted sentences “*X* are *A* and not *B*,” likewise involving several inclusion fallacies, as most probable.

The H11 scenarios yielded the strongest deviations from the predictions (to be discussed below).

For the other scenarios (H4, NEITHER NOR; H5, Affirmation *A*; H6, EITHER OR), the results corroborated both the predicted dominant selections and an overall high correspondence between BL and the data.

Figure 3 additionally provides strong evidence for a high similarity between results in the two conditions, the individual and general prediction tasks.

These overall results need supplementation from further measures, in order to assess results on the level of the single scenarios. Accumulation of the results of four scenarios of a type, such as in Figure 3, increases the *N* (reducing chance-findings) and excludes confounding factors specific to single tasks; but such results will tend to yield too positive a picture.

Table 1: Mean correlations between model-predictions and results, as well as between the two kinds of predication (individual and general) for the single scenarios in the six types H1 to H9) and a further noise class

	H1	H2	H11	H4	H5	H9	noise
$r_{B1\ B2}$	.85	.71	.79	.89	.92	.82	.75
$r_{B1\ M}$	.78	.67	.41	.63	.80	.72	.24
$r_{B2\ M}$	.78	.89	.43	.73	.91	.83	.34

Table 1, despite using averages, for this reason focuses on correlations for the *single* scenarios in a class. Please note, this differs from correlations on the accumulated level (which actually yield correlations .09 higher on average). Table 1 shows correlations between model and results, as well as between individual and general predication-conditions. The average correlations were all positive and generally large (or very large).

Only for H11 the average correlations with the model were only moderately positive. This was likewise the case for the high-noise scenarios (where predictions did not favor specific connectives). In these two classes, the number of tautology-selections was higher than expected (H15). The two deviations may be explained along the same lines: as mentioned, the OR-scenarios, just as the noise-scenarios, had much less clear predictions than all other scenario-classes. The second- and thirdmost frequently predicted hypotheses were not much less frequently predicted than the OR hypotheses themselves. Moreover, even for OR-predictions (based on specific participants of Pre-test 2), the second-highest pattern-probability, did not generally differ substantially from the second-highest. Such uncertainties in both may have led to the selections of H15, which suggested that everything is possible.

Even if the 30 scenarios were analysed individually (which cannot be done here) the overall pattern would remain similar. All 90 calculated correlations were positive, and only 14 % yielded $r < 40$ (particularly in the mentioned classes). The examples from Figure 2 with dominant AND, EITHER OR, A and OR predictions, for instance, corroborated the predicted dominant selections and had a high model-fit. Nonetheless, a low number of correlations did not show the overall positive results (even outside of the two mentioned classes), with values close to 0 and asymmetrical findings for both specific and general predications.

General Discussion

The findings corroborate that people do not judge probabilities extensionally, but instead allow for exceptions. Participants systematically committed a large number of inclusion fallacies (generalizing CFs, cf. von Sydow, 2009; von Sydow & Fiedler, 2012). Pattern probabilities, as formalized by BL, were shown to be quite successful in modeling the probability judgments in a multitude of scenarios only indirectly based on frequency estimates in Preliminary Study 2. Other models of CF have not yet been explicitly designed to test for these connectives, but it is as yet highly implausible that some adaptation of these models (e.g., confirmation, inverse probability, representativeness, averaging, quantum logic, support theory, rescaling, signed summation, etc.) will easily account for these data equally well without adopting the very idea of pattern probabilities themselves (cf. von Sydow, 2009, 2011, for more details).

Moreover, there was a large similarity between probability judgments about general predication and prediction about singular subjects. Combining BL with the auxiliary hypothesis that people may base judgments about singular nouns on hypothetical sampling led to quite successful predictions. However, it needs to be mentioned that the current finding might be limited to a generic interpretation of the singular subjects (e.g., "a professional basketball player (PBP) is/has"), although the introduction suggested an individual reading ("imagine one PBP"; "propositions about a PBP"). Further research is needed to investigate the role of different formulations in more detail.

Finally, some quantitative deviations from the predictions were found—some in a quite explainable manner regarding two classes of scenarios— but unsystematic deviations for single scenarios were found as well. Although this did not substantially alter the main findings, it may suggest that further factors are at work. To state it more emphatically, in my view, it would be a surprise if there were in fact no additional factors:

(1) Despite favoring BL as important account for CFs, I think there may be several other causes of CFs as well. I have mentioned other theories above. An example is that people in some contexts may reasonably be interested in the increase of probabilities (confirmation) instead of probabilities themselves (e.g., Lagnado & Shanks, 2002; Tentori, et al., 2012); and there may well be situations where people are interested in a synthesis of confirmation and the pattern idea: the degree of confirmation of different logical patterns.

(2) Even if focusing on pattern probabilities the current formalization of BL may only be one sub-class of modeling pattern probabilities in real-world predications. The current formalization is concerned with dyadic classes and assumes an equally weighted 2×2 input. Considerations to be examined include: (a) Although dyadic dichotomous logic as well as human language is often concerned with dichotomous (or dichotomized) categories, the implicit number of relevant categories can vary and may well matter. (b) The dichotomous classes need not refer to categorical classes, but can point to an underlying ordinal, interval, or rational scale. This may require a modified pattern approach that weights extreme cases more heavily. A domain specific approach to rationality that takes preconditions of models seriously should be sensitive to such aspects. (c) The categories and resulting frequency estimates may be defined in either an absolute (larger than 1.80) or a relative way (large). If the latter, BL's input may needs to be modelled depending on the context. (d) Within the present model, different contexts (scenarios) may lead to different noise-priors (reflecting the learned tolerance for exceptions for different scenario types), whereas here always flat noise-prior was used (cf. von Sydow, 2011). (e) People might assign important properties more weight than unimportant properties.

(3) The relationship between individual and general predication is presumably more intricate than assumed in the studies. As we have seen, BL is designed for general predication with a frequency-input based on a 2×2 contingency matrix. This assumes that evidence is ordered as cases to be assigned to one of the table's four cells. The results presented here support the idea that one can model individual probabilistic predication along the same lines by imagining 100 individuals and finding the highest pattern probability for statements as "A sophomore either owns a flat or shares a flat" (Fig. 2b). This seems unproblematic, since individual sophomores still fall into one of the four classes. The EITHER-OR here only expresses a lack of knowledge about which of two classes the individual is actually fits. Nevertheless, sentences such as "this ape from species *X* is either aggressive (*A*) or curious (*C*)" need not indicate lack of information, but rather an alternative meaning: that is, the individual ape may have been *A* (without being *C*) and at other times the reverse. Notably, the input-assumption would still hold for individual predication on a sub-individual event level, but on the group level this apparently positive extension of BL to individual predication now seems problematic; for, if an individual is "either *A* or *C*" it no longer fits any of the four classes (*A* & *C*, *A* & non-*C*, non-*A* & *C*, or non-*A* & non-*C*). This problem may be solved by adding ½ to both relevant cells, resulting at least in similar predications for both levels. Nonetheless, the issue remains problematic if we are concerned with heterogeneous groups of individuals, where in most cases *X* are either "*A* or *C* or both" or "*A* and *C*" (each based on sub-individual frequency information). The inclusive predicate "*A* or C or both" alone may be inappropriate. Participants may be interested in a pattern of patterns (*X* are $(A \wedge C) >< (A \vee C)$); and interpreted as pattern, this does not needs to be equivalent to $A >< C$, as actually valid in propositional logic. Such a pattern-of-pattern interpretation would not only be an interesting field of future research, but it might discourage the selection predicated by standard BL which assumes the absence of sub-classes.

In summary, real-life predication as well as probability judgments about these logical predications may plausibly be affected by a variety of additional factors, either external or ones calling for other more context-sensitive formalizations of pattern probabilities. In the light of such suggestions, the basic BL model was shown here to be surprisingly successful in accounting for a great variety of probability judgments about general as well as individual predications.

Acknowledgments

Aspects of this research, presented here in written form have previously been introduced in talks. I am grateful for the inspiring discussions at the Universities of Göttingen (M. R. Waldmann group) and Heidelberg (K. Fiedler group). I am also grateful for help from Martha Cunningham, Johanna Frisch and Dennis Hebbelmann and for the feedback of the anonymous reviewers. This work was supported by the grant Sy 111/2-1 from the Deutsche Forschungsgemeinschaft (DFG) as part of the priority program *New Frameworks of Rationality* (SPP 1516).

References

Chater, N., Tenenbaum, J., & Yuille, A. (2006). Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Science, 10*, 287-291.

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research, 50*, 123-129.

Gigerenzer, G. (1996). On narrow norms and vague heuristics. *Psychological Review, 103*, 592-596.

Hertwig, R., Benz, B., & Krauss, B. S. (2008). The conjunction fallacy and the many meanings of and. *Cognition, 108*, 740-753.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment, *Psych. Bulletin, 118*, 248-271.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S. & Caverni, J.-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106*, 62-88.

Kahneman, & Tversky (1996). On the Reality of Cognitive Illusions. *Psychological Review, 103*, 582-591.

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior, 36*, 210-226.

Lagnado, D. A. & Shanks, D. R. (2002). Probability judgment in hierarchical learning: A conflict between predictiveness and coherence. *Cognition, 93*, 81-112.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality. The probabilistic approach to human reasoning*. Oxford Univ. Press.

Schurz, G. (2005). Non-monotonic reasoning from an evolutionary viewpoint: ontic, logical and cognitive foundations. *Synthese, 146*, 37–51.

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions. *Organizational Behavior and Human Processes, 91*, 296-309.

Tentori, K., Crupi, V., & Russo, S. (2012). Determinants of the conjunction fallacy: Confirmation versus probability. *Journal of Experimental Psychology: General*.

Tenenbaum, J. & Griffith, Th. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences, 24 (4)*, 629-640.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293-315.

von Sydow, M. (2009). On a General Bayesian Pattern Logic of Frequency-Based Logical Inclusion Fallacies. *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society* (pp. 248-253). Austin, TX: Cognitive Science Society.

von Sydow, M. (2011). The Bayesian Logic of Frequency-Based Conjunction Fallacies. *Journal of Mathematical Psychology, 55(2)*, 119-139.

von Sydow, M. & Fiedler, K. (2012). Bayesian Logic and Trial-by-trial Learning. *Proceedings of the Thirty-Fourth Annual Conference of the Cognitive Science Society* (pp. 1090 - 1095). Austin, TX: Cognitive Science Society.