On a General Bayesian Pattern Logic of Frequency-Based Logical Inclusion Fallacies

Momme von Sydow (momme.von-sydow@bio.uni-goettingen.de)
Department of Psychology, Georg-August-Universität Göttingen, Gosslerstr. 14, D-37073 Göttingen, Germany

Abstract

Bayesian logic provides a rational model of probability judgments deviating from the standard extensional norm of extensional probability. It formalizes the general idea of an inductive pattern logic that may resolve paradoxes of inclusion. Bayesian logic predicts that it should be possible to generalize the phenomenon of frequency-based logical conjunction fallacies to a system of logical inclusion fallacies. In Experiment 1 quantitative conditions for conjunction fallacies and the role of negations are investigated. Experiment 2 provides a first test of the postulated more general system of logical inclusion fallacies. The results of both experiments confirmed the proposed pattern logic and its formalization as Bayesian logic. Other theories of the conjunction fallacy cannot readily explain this class of frequency-based and pattern-based inclusion fallacies. Whether there are simpler heuristics that may perhaps explain these data as well should be investigated in the future.

Keywords: Probability judgments; conjunction fallacy; logic; predication; paradoxes of inclusion.

Predication: Paradoxes of Logic and Probability Theory

Human predication about a type of entities (the subject class *X*) often deals with a logical relation (here using the general symbol: $\bigcirc_l$) between two predicates (predicate *A* and predicate *B*). Take the ubiquity of the predication of the relation '*A* and *B'*, understood as logical conjunction ($A \bigcirc_l B = A \wedge B$): "Ravens are black and they can fly." "Professional basketball players are quick and tall." Formulated generally: "*X* are *A* and *B"*.

Traditionally, a necessary condition for a valid logical predication is that the predicate is true, in the sense that *all* observed empirical cases confirm the classes of a connective's truth table that are true (Frege, Russell, Whitehead, Wittgenstein, Popper). Using predicate logic, one would treat a predication as involving an implicit universal quantifier, "*all X* are *A* and *B*". (An alternative is to assume an implicit existential quantifier, cf. Brentano's works.)

Such an interpretation of predication entails two problems: (1) The problem of exceptions: There are exceptions to most general predications that are concerned with contingent empirical matters (cf. the problem of monotonicity). Actually, there are for example white ravens (whole species of white ravens and cases of albinism) and there are ravens that cannot fly. (2) The problem of inclusions: The predication "ravens are black and they can fly" paradoxically entails that one may equally use the predicate of the inclusive disjunction "ravens are black *OR* they can fly or both" ($\forall x\ (B(x) \wedge F(x)) \Rightarrow (B(x) \vee F(x))$). Although this is, of course, true according to (extensional) logic, it is a psychologically uninformative conclusion, given that we are interested in an overall characterization of a situation. For *any* situation in which one would use an AND-predicate, one could equally use an OR-predicate. This problem may also lie at the heart of the paradoxes of implication in the narrow sense (cf. Evans & Over, 2004). From "ravens are black and they can fly" it follows that "If ravens are black then they can fly" ($(B \wedge F) \Rightarrow (B \rightarrow F)$). There is a bunch of these paradoxes, discussed in the literature, e.g.: $\neg B \Rightarrow (B \rightarrow F)$; $F \Rightarrow (B \rightarrow F)$; $(\neg B \wedge F) \Rightarrow (B \rightarrow F)$; $(\neg B \wedge \neg F) \Rightarrow (B \rightarrow F)$. Still, here we will not directly concentrate on conditionals, because there may be specific problems of conditionals, due to a subjunctive interpretation, an incomplete representation or a reference to causal relations. However, we saw that analogous paradoxes arise for other connectives as well. I generally call them paradoxes of inclusion. Even if one takes a different view on the paradoxes of implication, this does not resolve the paradox that $\forall x\ B(x) \wedge F(x)$ *implies* a justification for the predication $\forall R\ B(x) \vee F(x)$. The subset relation between an AND-set and an OR-set is generally undisputed. We here investigate this more general subset relation without investigating questions of implication directly.

An apparent solution to the first problem, the problem of exceptions, is the use of high probabilities as a necessary condition for prediction (apart from existential prediction). The sentence "Ravens are black and they can fly" may be interpreted as a claim that the probability $P(B\ \&\ F|R)$ is high (cf. Adams, 1986). This would offer a rational basis for predication, but would not resolve the paradox of inclusion.

Conjunction Fallacy Debate: Problems of Extensional Probability Judgments

A related debate – but normally discussed separately – is concerned with the assignment of probabilities of conjunctive events. Here subjects' intuition seems to deviate from the perhaps most simple norm of (extensional) probability theory (Kolmogorov, 1933):

In the classical Linda task, first proposed by Tversky and Kahneman (1982, 1983; cf. Kahneman & Frederick, 2002) participants were presented with a story: "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations." As dependent variable subjects had to rank statements to their probability, including: "Linda is a bank teller" (*B*) and "Linda is a bank teller and she's active in the feminist movement" (*B* and *F*). Most participants estimate $P(B \text{ and } F) > P(B)$. This is known as *conjunction fallacy (CF)*.

According to extensional probability theory, the conjunction '$B \wedge F$' can never be more probable than one of its components, '*B*' or '*F*' respectively, since any second property delimits the set size of the first one. Likewise also non-standard accounts of probability, like the Dempster-Shafer belief functions, Cohen's Baconian probabilities, or fuzzy logic, which partly gave up the strict additivity of probabilities, accept the conjunction rule: $P(A \wedge B) \leq P(A)$ (cf. Kahneman & Frederick, 2002). Tversky and Kahneman explained this phenomenon by the use of a "representativeness heuristic". This heuristic consists of a first step of evoking a prototype (a representative exemplar) to represent categories (e.g. bank teller) and a second step of judging the similarity of Linda with that prototypical exemplar (Kahneman & Frederick, 2002). Previously, the concept of representativeness has been criticized as being defined in a way that is too vague to count as a proper explanation (Wolford, 1991; Gigerenzer, 1996, cf. 2000). However, the mentioned refinement will not rule out this criticism completely.

A multitude of other qualitative and quantitative approaches have been inspired by this phenomenon (cf. v. Sydow, 2008, subm. a). Researchers argued qualitatively that the fallacy could be prevented if one used clear logical terms (e.g., Hilton, 1995; Mellers, Hertwig, & Kahneman, 2001; cf. Sides, Osherson, Bonini, & Viale, 2002; Bonini, Tentori, & Osherson, 2004), a frequency format (Fiedler, 1988; Gigerenzer, 1996, 2000; Hertwig & Gigerenzer, 1999), less complex tasks (Lagnado & Shanks, 2002), a ratings instead of rankings (Hertwig & Chase, 1998; Hertwig & Gigerenzer, 1999; Sloman et al., 2003; Wedell & Moro, 2008), or clear set inclusions (Sloman, Over, Slovak, & Stibel, 2003; cf. Johnson-Laird, Legrenzi, Girotto,& Legrenzi, 1999). Although these factors have shown to be important extensional cues, the pattern logic advocated here predicts that they are not sufficient to elicit extensional answers (v. Sydow, 2007, 2008). Actually, conjunction fallacies continued to occur even if one introduced many or all of these factors simultaneously (Lagnardo & Shanks, 2002; Tentori et al., 2004; Sides et al., 2002; Bonini et al., 2004; v. Sydow, 2007, 2008).

Moreover, there are a number of interesting quantitative accounts of CFs that may in principle provide predictions also for the frequency-based CFs investigated here. It has for instance been argued that the conditional probability of an AND-relation given the data, $P(A \;\&\; B|\; D)$, is misunderstood as inverse probability (cf. Wolford, 1991; Fisk, 1996; Hertwig & Chase, 1998), or as an increase in probability (support theory; Lagnado & Shanks, 2002; Sides et al., 2002; cf. Crupi, Fitelson, & Tentori, 2008) or as some other measure (Busemeyer & Wang, 2007; Costello, 2005). Although these models were able to explain some data, the empirical situation has remained unclear and none of these models would have predicted the frequency-based CFs confirming Bayesian Logic (v. Sydow, 2007, 2008).

Bayesian Pattern Logic

General Idea Bayesian Logic (v. Sydow, 2007, 2008, subm. a, subm. b), formalizes the idea of pattern-based, logical probability judgments, providing a rational basis for a class of conjunction fallacies and related logical fallacies. This account builds on the idea that probability is a polysemous notion (Hertwig & Gigerenzer, 1998; cf. Fiedler, 1988; Wolford, 1991; Fisk, 1996; Sides et al., 2002). However, from my point of view this does not need to imply normative agnosticism (Kahneman & Tversky, 1996). Extensional probabilities, e.g. $P_E(A \wedge B|D)$, are defined by (the limit of) the number of confirmatory cases (the extension) relative to all observed cases in the given universe of discourse *X*. The analogous pattern probability ($P_P(A \wedge B|D)$), specified by Bayesian logic, is the probability that, given some data, the underlying *generative* structure as a *whole* is e.g. an AND probability *pattern*.

There are two main reasons for introducing pattern probabilities. First, pattern probabilities could provide a rational probabilistic basis for logical predication that can account for exceptions and that may resolve the paradoxes of inclusion. Bayesian logic allows using a high probability criterion as necessary condition for predication, but does not fall pray to the paradoxes of inclusion. "Ravens are black and they can fly" is probable, "Ravens are black or they can fly or both or none" may nonetheless be improbable (as pattern). Second, the goal of probability judgments is often to characterise an overall situation in one single sentence. Imagine a hunter, who has knowledge about a few exemplars of a species of wisent and the relation of two properties (say thick fur, *A,* and good meat, *B*). Think of two worlds (Table 1).

Table 1: Observed frequencies in two worlds W1 and W2

W1	*B*	¬*B*	W2	*B*	¬*B*
A	18	5	*A*	18	15
¬*A*	6	7	¬*A*	1	2

In which of these worlds would the hunter tell his friend that probably wisent have thick fur and at the same time provide good meat? The extensional probability $P_E(A \wedge B)$ is, of course, identical in both worlds. The pattern probability favours a dominant AND-pattern in world 1 and an A-pattern in world 2 (v. Sydow, subm. a). If we really have the goal to use probability judgments to evaluate overall situations, extensional probabilities are inefficient. An AND-proposition would only be concerned with *one* cell relative to all other cells, without characterising the overall pattern. You would need several extensional probability judgments to characterise a situation. If people have the goal to characterise a situation in one sentence, a rational analysis (cf. Anderson, 1990) would predict that people should intuitively reason in a pattern way, deviating from extensional probability judgments.

Sketch of the Model Bayesian (Pattern) Logic (or BL, for short) is intended to be a rational and, for the time being, an empirical model of pattern probabilities (for more details cf. v. Sydow, 2007, 2008, subm. a, b). However, the model is formulated on the so-called computational, not on the algorithmic level (sensu Marr). It is assumed that people are intuitive experts of pattern recognition and approximate the predictions of BL, whatever underlying process gives rise to this performance.

BL aims to contribute to the renaissance of Bayesian models in the cognitive sciences (cf. Oaksford & Chater, 2007; Tenenbaum, Griffith & Kemp, 2006). BL provides probabilities about hypotheses concerned with logical relations and whole situations (previous Bayesian accounts of the CF were not, cf. Fisk, 1996). BL itself is formulated in terms of extensional probability theory, but on the emergent level, it does not subscribe to extensionality and the conjunction rule. BL calculates pattern-sensitive posterior probabilities of logical hypotheses, given some data pattern D in a 2×2 contingency matrix: $P_P(A \bigcirc_l B \mid D)$.

The 'logical' hypotheses are formalized as probabilistic probability tables, consisting of four probabilities. The changing posteriors are actually second order probabilities each concerned with a whole probability table.

(1) Probability tables (PTs) are built corresponding to the 16 truth tables of propositional logic. The construction of the PTs relies on two assumptions.

Assumption of idealization. To get from truth tables to PTs an assumption about the distribution of the probabilities of the true cases has to be made. Formal logic only assumes that evidences correspond to true cases. According to the assumption of idealization the cell probabilities of true cases, $P_{PT}(T)$, in a PT are assumed to be equally probable (cf. Johnson-Laird et al., 1999). This formalizes the idea of an ideal logical explanation. In case of deterministic logical relationships (no noise, $r = 0$), the cell probability of a true cell in an ideal logical PT is one divided by the number of true cells: $P_{PT}(T \mid r = 0) = 1 / N(T_l) = t_l$, resulting for instance in $t_{A\wedge B} = 1$, $t_A = 50$, and $t_{A\vee B} = .33$. For deterministic relations the probability of a false cell is zero: $P_{PT}(F \mid r = 0) = 0$.

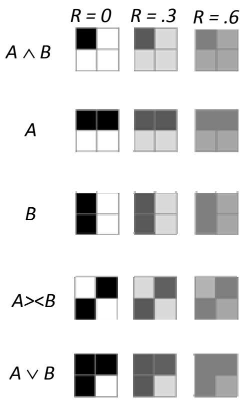


Figure 1: Illustration of the general idea of different PTs and noise-levels r. Darker vs. brighter cells represent high vs. low probability.

Secondly, the *assumption of uncertainty* formalizes the idea of *probabilistic* logical relations. Here a general level of uncertainty for a PT is assumed (cf. Oaksford & Chater, 2007), given that no specific sources of noise are known (v. Sydow, 2007, 2008). Whereas a single disconfirmatory event should be able to falsify a deterministic relationship, this event should not refute a probabilistic relationship. A single PT formalizes an explanatory connective at a particular noise level. The assumed noise level r of a PT is taken to be the ratio of the cell probability of a true cell relative the probability of a false cell.[1] The same formalization results if noise is taken to be an external cause that influences each cell independently (noisy-or). Based on this assumptions, it can be shown that for ideal noisy PTs the probability of a true cell is $P(T) = 1 / (t + 4r - tr)$ and that of a false cell is $P(F) = r / (t + 4r - tr)$.

For $r = 0.1$ the PT for $A \wedge B$ is (.76, .08, .08, .08), the PT for $A\vee B$ is (.32; .32; .32; .04). Each PT represents a possible connective-noise combination: $A \bigcirc_{l,r} B$. Note, that here $P(r)$ is not treated as free parameter. We start with a flat uninformative prior probability distribution both for the noise levels and the connectives.

[1] Previously r was the proportion of probabilities that were distributed equally over all cells of a PT (cf. v. Sydow, 2007, 2008). Both possible variants here lead to the same predictions.

(2) Further model steps. Now it is possible to generate the likelihood for the generative connective-noise hypotheses, $A \bigcirc_{l,r} B$. The multinomial distribution provides the discrete probability distribution of obtaining a data pattern D of the disjoint outcomes $D = (x_a, x_b, x_c, x_d)$, in a sample of N trials, given a PT (a connective-noise hypothesis) with the respective probabilities p_a, p_b, p_c, p_d (with $0 \le p_m \le 1$).

$$P_H(D \mid A \circ_{l,r} B) = \binom{N}{x_a\, x_b\, x_c\, x_d} p_a^{x_a}\, p_b^{x_b}\, p_c^{x_c}\, p_d^{x_d} \quad (1)$$

Subsequently, Bayes' theorem is used to calculate the posterior probabilities given the data, D:

$$P_H(A \circ_{l,r} B \mid D) = \frac{P(D \mid A \circ_{l,r} B)P(A \circ_{l,r} B)}{P(D)} \quad (2)$$

$$P(D) = \sum_l \sum_r P(D \mid A \circ_{l,r} B)P(A \circ_{l,r} B) \quad (3)$$

In order to calculate the probability of a *logical* hypothesis ($H_l = A \bigcirc_l B$), we sum up the posterior probabilities over all error levels r_j:

$$P_H(A \circ_l B \mid D) = \sum_j P_H(A \circ_{l,r} B \mid D) \quad (4)$$

The resulting value is a pattern probability, with the same form as the requested conditional probability $P(A \circ_l B|D)$. Since subjects may be not interested in differences, but in ratios of these probabilities one may use the mean of the log odds of all comparisons between hypotheses (cf. Study 1).

(3) *Predictions and previous corroborations*. BL makes a number of novel predictions. BL should be applicable if one is concerned with alternative logical hypotheses and whole situations (rather than subsets). Qualitatively, CFs should hence occur, even if one uses extensional cues, as long these preconditions of the model apply. In von Sydow (2008) this was corroborated, using frequencies, clear set inclusions and rating scales. Quantitatively, BL predicts *sample size effects* (v. Sydow, 2007) and *pattern sensitivity* (v. Sydow, 2008).

The following two experiments test some further predictions. Experiment 1 is concerned with the role of negations and quantitative conditions of double conjunction fallacies. Experiment 2 is concerned with the generalization to other connectives: logical inclusion fallacies in general.

Experiment 1: Quantitative Conditions of Double Conjunction Fallacies and Negations

Goals Experiment 1 had three goals: First, the quantitative conditions of double CFs will be investigated. Second, the experiment was intended to check whether the predictions hold independently of whether we use negated evidence and negated hypotheses. Third, the experiment excluded possible misunderstandings of the conjunction and the affirmation by using all four kinds of possible conjunctions.

Double CFs are committed if participants probability judgments involve $P_{emp}(A) < P_{emp}(A \wedge B)$ and $P_{emp}(A \wedge B) > P_{emp}(B)$ (cf. Carlson & Yates, 1989). Bayesian logic formalizes the proposed pattern logic and provides more precise predictions when these double CFs are rationally predicted. In this experiment all four types of data patterns had a modal (conjunctive) frequency in the same logical cell (cf. Fig. 2A first line). However, dominant double CFs are only predicted for two of these patterns (Pattern type 1 and 2; cf. Fig. 1B). Pattern type 1 investigated whether dominant double CFs occur at least with highly probable conjunctive cell (Pattern Type 1 with $P_E(B \wedge F) = .8$). Pattern type 2 explores whether participants commit double CFs even if its probability is below 50 % ($P_E(B \wedge F) = .4$). BL predicts dominant double CFs with reduced confidence, cf. Fig. 1B). This would pose a problem for an account that is only concerned with very high probabilities (cf. Adams, 1986). For the other two patterns, double CFs should not be dominant, although the extensional probability of the conjunction is now higher (Pattern Type 3 with $P_E(B \wedge F) > .5$) or gets closer to an AND-pattern (Pattern Type 4). The few theories that explicitly predict double CFs (Costello, 2005, cf. Carlson & Yates, 1989) would have to predict them here as well. However, Bayesian logic predicts more double CFs in Pattern 2 than in Pattern 3.

Method 64 participants from the University of Göttingen participated in the study voluntarily. Participants received a booklet with four judgment tasks, each presented on a single page. The shown frequencies were derived from four pattern types, using 4 rotations as balancing factors, resulting in 16 frequency patterns (Fig. 2A). The sample size was always 60. Participants received all four pattern types, but only one rotation. Each pattern type and rotation occurred equally often in all 16 presentation positions (cf. Fig. 2C).

Participants were instructed to be in a role of someone who is generally interested in the overall evaluation of different schools (e.g. the ‘Linda school’). They should tick the hypothesis about the logical relation between two properties of the graduates (they become bank teller, they become feminist) that is “at least most probable”.

The frequency information was presented to the participants in a contingency table, whose cells and marginals were labeled (e.g. ‘bank teller & feminist’, ‘bank teller’) The used eight hypotheses read: “The graduates of the Linda School generally became bank tellers and at the same time feminists” (*B* & *F*) and analogous formulations for *B* & *Non-F; Non-B* & *F*; *Non-B* & *Non-F*. The affirmations read “[…] bank tellers” (*B*); “[…] no bank tellers” (*non-B*), with analogous other formulations for *F* and *non-F*. Subjects were asked to answer intuitively.

Results Figure 1D shows the percentage of the selection of a particular logical hypothesis to be most probable in the four pattern types. The results were conflated over the factors Rotation and Position. Beforehand, the data for the rotations were recoded (e.g., *B* & *non-F* selections in the 90° rotations became *B* & *F* selections). In all pattern types the dominant selections corresponded to the dominant prediction of BL. As expected, the results for the Pattern types, where the confidence should be lower, were slightly less pronounced. The second most frequent selections for all pattern types corresponded to the second most probable pattern, as predicted by BL (cf. Fig. 2B). However, this may correspond to a combination of other strategies (extensional strategies in type 1 and 2, and a most probable cell strategy in pattern 3 and 4).

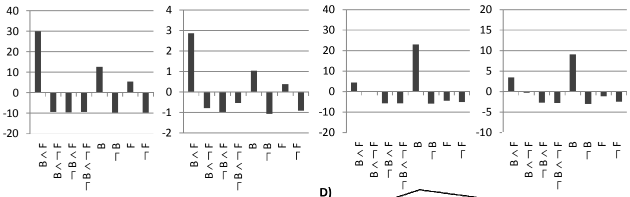
If participants misunderstood conjunctions as disjunctions, they should have answered differently.

A)

	Pattern type 1			Pattern type 2			Pattern type 3			Pattern type 4		
Rotation 0°	1a Linda school			2a Sara school			3a Maria school			4a Nina school		
		F	¬*F*		*F*	¬*F*		*F*	¬*F*		*F*	¬*F*
	B	48	6	*B*	24	12	*B*	31	26	*B*	27	21
	¬*B*	2	4	¬*B*	10	14	¬*B*	1	2	¬*B*	7	5
Rotation 90°	1b Lena school			2b Sabina school			3b Martina school			4b Nele school		
		F	¬*F*		*F*	¬*F*		*F*	¬*F*		*F*	¬*F*
	B	2	48	*B*	10	24	*B*	1	31	*B*	7	27
	¬*B*	4	6	¬*B*	14	12	¬*B*	2	26	¬*B*	5	21

Rotation 180° (Leonie, Steffi, Marlene, Nadine) and rotation 270° (Lara, Selma, Miriam, Nora) is not shown.

B)



C)

Order	t1	t2	t3	t4
Order 1	1a	2b	3c	4d
Order 2	1c	2a	3d	4b
Order 3	1b	2d	3a	4c
Order 4	1d	2c	3b	4a
Order 5	2a	4b	1c	3d
Order 6	2c	4a	1d	3b
Order 7	2b	4d	1a	3c
Order 8	2d	4c	1b	3a

Order	t1	t2	t3	t4
Order 9	3a	1b	4c	2d
Order 10	3c	1a	4d	2b
Order 11	3b	1d	4a	2c
Order 12	3d	1c	4b	2a
Order 13	4a	3b	2c	1d
Order 14	4c	3a	2d	1b
Order 15	4b	3d	2a	1c
Order 16	4d	3c	2b	1a

D)

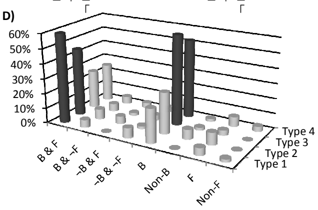


Figure 2: (A) Shown data patterns. (B) Predictions (ordinate: mean log odds of posterior probabilities). (C) 16 presentation orders for 4 positions: numbers = pattern types, letters = rotations. (D) Percentage of selections for types (recoded).

The few other theories that explicitly predicted double CFs would have predicted them in *all four* pattern types (cf. Carlson & Yates, 1989; Costello, 2005). Other theories would predict specific order effects.

A parsimonious way to test, whether only pattern type or also other factors contributed to the data is to conduct a log-linear analysis, simply treating the data as independent observations. On the level of single data points, Pattern type and Rotation are fully crossed factors and Position can be used as a third factor, not considering confounding predecessors (order). To increase the number of expected cases in each cell we used the factors Pattern types (Types 1 & 2 vs. Types 3 & 4) * Rotations * Positions (Position 1 & 2 vs. 3 & 4) * possible answers (*B* & *F* vs. *B* vs. Rest). If the answers are only determined by the factor postulated by BL, a hierarchical model with only one interaction between Pattern type and answers shall fit the data. Our analysis yielded that only the 2-way term for the interaction became significant ($k = 2$ with Pearson’s $\chi^2 = 31.26$, $df = 17$, $p = .01$). The default model fitting then eliminated all two-way term interactions that did not significantly contribute to this effect, resulting only in the predicted second order term, which fitted the data without significant deviation ($\chi^2 = 28.60$, $df = 42$, $p = .94$). A further analysis was run, enlarging n per cell by distinguishing only two types of rotations, again yielding the same best model with only the predicted second order term (Pattern type * Answers) ($\chi^2 = 12.12$, $df = 18$, $p = .84$). As predicted, participants’ selection of the most probable hypothesis was primarily affected by Pattern type, not by Rotation or by Position.

Discussion The results of Experiment 1 provide evidence that double CFs occur systematically and under the quantitative conditions predicted by BL. The results corroborated the existence of a kind of pattern logic of probability judgments and the postulated metric of BL (e.g. comparing Type 2 and Type 3). The used eight hypotheses with the frequency variations excluded a misunderstanding of the ‘*B* AND *F’* as ‘*B* OR *F’*, and of the ‘*B’* as ‘*B* & *non-F’*. It was shown that in the tasks without memory component BL holds for negations as well as for affirmations.

Experiment 2: On a Pattern Logic of Logical Inclusion Fallacies

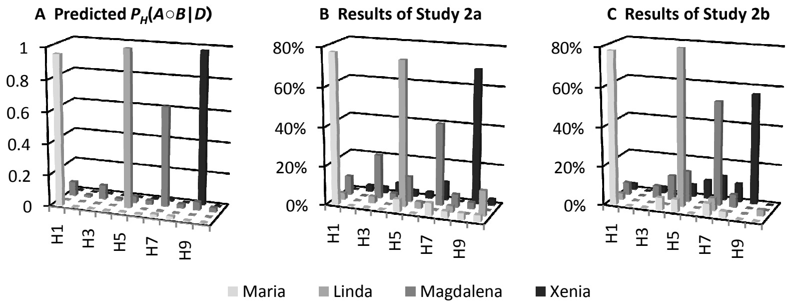
Experiment 2a and 2b were aimed to investigate logical inclusion fallacies more generally. Whereas previous research has concentrated on conjunctions and affirmations, and only a few contributions discussed disjunctions (cf. Carlson & Yates, 1989; Costello, 2005), we will investigate *several* logical connectives at the same time here (without any filler hypothesis). Bayesian logic predicts a system of logical inclusion fallacies and a resulting combination of inclusion fallacies like: $P(B) > P(B \wedge F) > P(B \vee F)$ or $P(B \wedge \neg F) > P(B >< F) > P(B \vee \neg F)$. Study 2b was mainly used to replicate the findings of Study 1a under stricter instructions and a different formulation for the inclusive disjunctions.

Method 64 students from the University of Göttingen (32 in each study) voluntarily participated. Participants received a booklet with four probability judgment tasks corresponding to a Maria, Linda, Magda, and Xenia school, each with frequency information about a sample of graduates. We used four presentation orders (MLM2X, LXMM2, XM2LM, M2MXL).

The task was about different girl’s schools and the probability of different hypotheses concerning the career and political attitude of their graduates. Participants should tick the hypothesis that they think is most probable, given the situation. The frequencies were represented in a contingency table, here using symbols (smileys) to represent number of observed instances. Again both the cells (e.g., Bank teller & active feminist) and the marginals (e.g. Bank teller) were labeled to prevent any confusion concerning set inclusions (cf. Sloman et al., 2003). The number of cases is reported in the caption of Figure 3. The AND-hypothesis, for instance, read “The graduates of this school generally become bank tellers AND at the same time active feminists” (Fig. 3D).

In Study 2a we explicitly used the sentence “You are not interested in particular subgroups, but in a different hypothesis (*H*) investigating the overall situation at these schools”. In Study 2b this was skipped to test whether pattern probabilities are even elicited without explicit reference to an interest in the overall situation. Additionally, we varied the formulation of the disjunction, using “become feminist bank tellers, OR ALSO non-feminist bank tellers, OR also feminist non-bank tellers” to exclude any misunderstanding about the disjunction hypothesis.

Main Results and Discussion The main results of the experiments are presented in Fig. 3B and C. It is apparent



D Hypotheses and Predictions

Hypotheses “The graduates of this school generally become…”	Maria	Linda	Magda	Xenia
(H1) “…**bank tellers**, whether they are active feminists or not” (*B*)	BL, $SU_{2,4}$	SR_2		
(H2) “…**no bank tellers**, whether they are active feminists or not” ($\neg B$)			IN_4, SU_4, $SR_{3,4}$	IN_4, $SR_{3,4}$
(H3) **“…active feminists**, whether they are bank tellers or not” (*F*)		IN_2, SR_2	IN_3, SU_3	
(H4) “…**no active feminists**, whether they are active feminists or not” ($\neg F$)				
(H5) **“…bank tellers AND at the same time active feminists”** ($B \wedge F$)	IN_2, SU_2	BL, IN, $SU_{3,4}$, SR	IN_3, SR_2	
(H6) “…**bank tellers AND at the same time no active feminists**” ($B \wedge \neg F$)	IN_4, IN_3, $SU_{3,4}$, SR		IN_2	IN_2, SR_2
(H7) “…**bank tellers OR also active feminists or also both**” ($B \vee F$)	EP		BL, EP	EP
(H8) **“…bank tellers OR also no active feminists or both**” ($B \vee F$)	EP, SU_2			
(H9) “…**EITHER bank tellers,** who are no active feminists, **OR active feminists**, who are no bank tellers” ($B >< F$)			IN_2, $SU_{2,4}$	BL, $IN_{2,3}$, SU, SR_2
(H10) “**IF** the graduates of this school become **bank tellers, THEN they become feminists**” ($B \rightarrow F$)		EP, SU_2, IN_2	SU_3, IN_3	

Figure 3: (A) Posterior probabilities of BL calculated for the Maria school ($x_a = 9$, $x_b = 9$, $x_c = 1$, $x_d = 1$), Linda school (16, 1, 1, 2); Magdalena school (7, 6, 6, 1), Xenia School (1, 9, 9, 1). (B, C) Percentage of selections of most probable hypotheses in Experiment 2a and 2b. (D) Formulations of the used hypotheses (H1 to H10) and modal predictions of Bayesian logic (BL), extensional probability (EX), inverse probability (IN), support - difference measure (SU), support - ratio measure (SR) (indexes refer to presentation positions specific to the four used presentation orders)

that the majority of answers corresponded to the modal predictions of Bayesian logic (cf. Fig. 3A). For each school the predicted answer was clearly above the chance level of $p_{\text{chance}} = 0.10$ (Maria: $\chi^2(1, n = 32) = 136.1$, $p < 0.0001$; Linda: $\chi^2(1, n = 32) = 150.2$, $p < 0.0001$; Magda: $\chi^2(1, n = 32) = 33.3$, $p < 0.0001$; Xenia: $\chi^2(1, n = 32) = 122.7$, $p < 0.0001$). We also derived predictions for extensional probability and influential alternative accounts of the conjunction fallacy, like inverse probability or support (Fig. 3D). Based on the presented frequencies, the latter accounts often falsely predicted $P_{emp}(B \text{ AND } F) > P_{emp}(B)$ in Maria school patterns as well (given experience $P_E(B \text{ and } F) < P_E(B \text{ and non-}F)$).

The results provide first evidence for a systematic of logical inclusion fallacies. The results even remained stable in Study 2b were we did not explicitly mention an interest in an overall evaluation. In such contexts the use of pattern probabilities even seems to be the default strategy.

Summary

A pattern logic may resolve paradoxes of predication and paradoxes of probability judgments. Bayesian logic is presented as a rational formalization of this pattern logic of probability judgments (cf. v. Sydow, 2007, 2008). The empirical results corroborated two new predictions of Bayesian logic. Experiment 1 confirmed the quantitative conditions under which Bayesian logic predicts double CFs.

Moreover, in a context with given frequencies and without memory load, we found no asymmetry in the treatment of positive and negative hypotheses and evidences. In Experiment 2 a generalization to other logical connectives was tested, corroborating the predicted system of resulting logical inclusion fallacies. The results additionally suggest that pattern logic may be a default strategy (Exp. 2a vs. 2b).

The found system of pattern-based probability judgments cannot be explained by any previous theory of the CF, like inverse probability, support, signed summation, or rescaling. Although BL has some similarities to representativeness, the results cannot be explained by a 'prototypic exemplar + similarity' explanation (cf. v. Sydow, 2007).

In conclusion, there seems to be a highly coherent class of frequency-based inclusion-fallacies, which are consistent with the rational predictions of Bayesian logic.

Acknowledgements

I would like to thank Johanna Frisch and Deborah Wolff for running the studies. I owe thanks to members of the CogSci lab at Göttingen University and to several anonymous reviewers for their feedback and advice. This research was supported by the DFG (grant 'Bayeslogik', Sy 111/1-1).

References

Adams, E. W. (1986). On the logic of high probability. *Journal of Philosophical Logic, 1*, 255-279.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Laurence Erlbaum Associates.

Bonini, N., Tentori, K., & Osherson, D. (2004). A different conjunction fallacy. *Mind & Language, 19*, 199-210.

Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision-making. *Journal of Mathematical Psychology, 50*, 220-241.

Carlson, B. W., & Yates, J. F. (1989). Disjunction errors in qualitative likelihood judgments. *Organizational Behavior and Human Decision Processes, 44,* 368-379.

Costello, F. J. (2005). A unified account of conjunction and disjunction fallacies in people's judgments of likelihood. In *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (pp. 494-499). Mahwah, NJ: Laurence Erlbaum Associates.

Crupi, V., Fitelson, B., & Tentori, K. (2008). Probability, confirmation, and the conjunction fallacy. *Thinking & Reasoning, 14*, 182–199.

Evans, St. B. T., & Over, D. E. (2004). *If*. Oxford University Press.

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

Fisk, J. E. (1996). The conjunction effect: Fallacy or Bayesian inference? *Organizational Behavior and Human Decision Processes, 67*, 76-90.

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. Oxford University Press.

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

Hertwig, R., & Gigerenzer, G. (1999). The 'Conjunction fallacy' revisited. How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12,* 275-305.

Hertwig, R., & Chase, V. M. (1998). Many reasons or just one: How response mode affects reasoning in the conjunction problem. *Thinking and Reasoning, 4*, 319-352.

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

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

Kahneman, D., & Frederick, S. (2002). Representativeness revised: Attribute substitution in intuitive judgment. In: T. Gilovich, D. Griffin, & D. Kahneman (Eds.). *Heurstics and biases.* Cambridge University Press.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*, 582-591.

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

Mellers, B. A., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science, 12*, 269–275.

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

Sides, A., Osherson, D., Bonini, N., & Viale, R. (2002). On the reality of the conjunction fallacy. *Memory and Cognition, 30*, 191-198.

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

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Science, 10*, 309-318.

Tentori, K., Bonini, N., & Osherson, D. (2004). The conjunction fallacy: A misunderstanding about conjunction? *Cognitive Science, 28*, 467-477.

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

von Sydow, M. (2008). The Bayesian logic of conjunction fallacies: Probability rating tasks and pattern-sensitivity. *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society* (pp. 1795-1800). Austin, TX.

von Sydow, M. (2007). The Bayesian logic of the conjunction fallacy: Effects of probabilities and frequencies in contingency tables. *Proceedings of the Twenty-Nineth Annual Conference of the Cognitive Science Society* (pp. 1611-1616). Austin, TX.

Wolford, G. (1991). The conjunction fallacy? A reply to Bar-Hillel. *Memory & Cognition, 19,* 415-417.

Wedell, D. H., & Moro, R. (2008). Testing boundary conditions for the conjunction fallacy: Effects of response mode, conceptual focus, and problem type. *Cognition, 107*, 105-136.