Running head: PATTERN-BASED LOGIC OF INCLUSION FALLACIES

Towards a Pattern-Based Logic of Probability Judgements and Logical Inclusion “Fallacies”

Momme von Sydow

Department of Psychology and IWR, University of Heidelberg, Germany,

MCMP, University of München, Germany

‘The Version of Record of this manuscript has been published
and is available In

Thinking & Reasoning (2016, online; in press),

DOI: 10.1080/13546783.2016.1140678

http://www.tandfonline.com/

Correspondence concerning this article should be addressed to:

Momme von Sydow, Department of Psychology,

University of Heidelberg, Hauptstr. 47, 69117 Heidelberg, Germany.

Email: Momme.von-Sydow@psychologie.uni-heidelberg.de,

Momme.von-Sydow@lrz.uni-muenchen.de

Abstract

Probability judgments entail a conjunction fallacy (CF) if a conjunction is estimated to be more probable than one of its conjuncts. In the context of predication of alternative logical hypothesis, Bayesian logic (BL) provides a formalization of pattern probabilities that renders a class of pattern-based CFs rational. BL predicts a complete system of other logical inclusion fallacies (IFs). A first test of this prediction is investigated here, using transparent tasks with clear set-inclusions, varying in observed frequencies only. Experiment 1 uses data where BL makes dominant predictions; Experiment 2's predictions were less clear, and we additionally investigated judgments about second-most probable hypotheses. The results corroborated a pattern-probability account and cannot be easily explained by other theories of CFs (e.g., inverse probability, confirmation). IFs were not limited to conjunctions, but rather occurred systematically for several logical connectives. Thus pattern-based probability judgements about logical relations may constitute a basic class of intuitive but potentially rational probability judgements.

Keywords:

Probability judgements; conjunction fallacy; inclusion fallacies; Bayesian models; rationality; hypothesis testing; inductive logic

Although intuitive probability judgements appear essential for an adequate understanding of an uncertain world, there is strong evidence that human probability judgements deviate substantially from the standard norms of probability theory (Gilovich, Griffin, & Kahneman, 2002; Nickerson, 2004; Fiedler & Juslin, 2006).

Conjunction fallacies (CF) seem to be particularly basic deviations from this norm; and these became a cornerstone of Tversky and Kahneman's "heuristic and biases" research program (Tversky & Kahneman, 1974, 1983; Kahneman & Frederick, 2002, 2005). They found that people committed a fallacy in judging conjunctions ('*A* and *B*') to be more probable than one of its conjuncts (e.g., '*A*') although this violates "[p]erhaps the simplest and most fundamental qualitative law of probability" (Tversky & Kahneman, 1983, p. 293), the conjunction rule, $P(A \wedge B) \leq P(A)$. Since we will here investigate several logical connectives, we will investigate a generalization of the conjunction rule: if event *E1* is included by event *E2* ($E1 \subseteq E2$), then $P(E1) \leq P(E2)$ holds (inclusion rule). Deviations from this will be called inclusion fallacies (IFs).

Although the bias-and-heuristic research program has been criticized for its failure to specify clearly defined algorithms explaining fallacies (Gigerenzer, 1996, 1998; Nilsson, Juslin, & Olsson, 2008), it cannot be doubted that the research questions have remained highly fertile (Fiedler & von Sydow, 2015). In particular, the debate on CFs continues to generate innovative ideas on the (ir)rationality of human belief (e.g., Busemeyer, Pothos, Franco, & Trueblood, 2011; Costello & Mathison, 2014; Crisp & Feeney, 2009; Gigerenzer, 2000; Hartmann & Meijs, 2012; Hertwig, Benz, & Krauss, 2008; Jarvstad & Hahn, 2011; Jenny, Rieskamp, & Nilsson, 2014; Lagnado & Shanks, 2002; Mellers, Hertwig, & Kahneman, 2001; Neace, Michaud, Bolling, Deer, & Zecevic, 2008; Schupbach, 2009; Sloman, Over, Slovak, & Stibel, 2003; Tentori, Crupi, & Russo, 2013; von Sydow, 2011; Wedell & Moro, 2008).

In the following, it is first argued that neither standard logic nor standard (extensional) probability can reasonably be applied directly as an adequacy criterion for logical predication, if logical predication is based on a fit between data and logical hypotheses. An alternative (intensional) interpretation of probability, supplementary to standard extensional probability, is proposed as a posterior probability that a specific generative logical rule holds, given the data. Second, 'Bayes logic' (BL) is presented as a possible mathematical formalisation of such a probability, systematically predicting not only CFs, but also other inclusion fallacies (IFs). . Third, we discuss influential alternative accounts of CFs that may be generalised to other IFs as well. Although these theories may provide reasonable explanations of other classes of CFs, we try to show that BL is the best model for the investigated class of CFs. Fourth, two experiments are reported investigating the predicted class of CFs and, particularly, a generalised system of IFs. Finally, the General Discussion addresses potential implications of the findings for BL and other theories, open questions, and whether we are here concerned with 'rational CFs'.

Bayes Logic: a Pattern-Based Model of Probability Judgement

Describing Logical Relationships: Extensional Probabilities vs. Intensional Pattern Probabilities

Building on the approach of rational analysis (Anderson, 1990; Chater & Oaksford, 2000) and the idea that one should not apply the 'narrow norm' of standard probability in a content-blind way, since 'probability' is polysemous (Hertwig & Gigerenzer, 1999; Gigerenzer, 1996; Teigen, 1994; cf. Hertwig & Volz, 2013), we argue here that some deviations from standard 'extensional' probability may be explainable by another type of probability (here called 'intensional' probability).

The ‘extension’ of a concept is given by the elements that fall under a given concept (designation), whereas the ‘intension’ of a concept is its ‘meaning’. Following Tversky & Kahneman (1983), I call probability measures here ‘extensional’ if and only if they fulfil the inclusion rule. Typically, this holds for the frequentist probability interpretation referring to the relative frequency of confirmatory elements (the extension) or its limit (Hájek, 2001; Gigerenzer, 1991; 2000).[1]

The contrastive notion ‘intension’ actually has a broader usage than covered by this paper. However, the ‘intensional probabilities’ formalised here refer to one central aspect of intension: they not only refer to the number of elements in a set (the extension), but also to the meaning (the intension) as far as it relates to the acceptance region of a logical proposition, and the *distribution* of elements inside or outside of region. Such intensional probabilities can be more specifically called logical ‘pattern probabilities’ (von Sydow, 2011; cf. von Sydow, 2009). This measure resembles the similarity, accuracy of fit, or minimal distance between an actual pattern of belief after obtaining data, and ideal explanatory logical patterns allowing for exceptions. We nay also call this ‘truthlikeness’.

Insert Table 1 about here

Table 1 provides an overview of the postulated extensional and intensional usage of probability. In a rational analysis (Anderson, 1990; Chater & Oaksford, 2000) it is basic to

[1] The inclusion rule likewise holds for most non-standard accounts of probability (cf. Hájek, 2001; Fitelson, 2006; Foley, 2009; Kern-Isberner, 2001; Schurz, 2005; cf. von Sydow, 2011).

specify a cognitive system's goals. Extensional and intensional probabilities answer different kinds of questions. Extensional probabilities are normative answers to a question such as: "How likely is it that an individual from a population is in a subset *H*, given the data?" In contrast, pattern probabilities answer the question: "How likely is it that an underlying (probabilistic) rule *H* operates, given the data?" I will argue that only the latter question can be reasonably applied to general predications about individuals or classes.

In this paper we are mainly interested in the predication of *dyadic* logical rules concerning *classes*. We are dealing with the probabilistic adequacy of propositions such as "*X* are *A* ○ *B*", with '○' signifying standard dyadic logical connectives, with '*A'* and '*B'* being dichotomic events (*A* vs. non-*A*, *B* vs. non-*B*). Leaving aside some problems regarding the correspondence between *ideal* language-connectives and their *ordinary* linguistic interpretations (Hertwig et al., 2008; von Sydow, 2014, 2015), general dyadic predication deals with alternative hypotheses *H*, such as: people of tribe *X* are "*A* AND *B*", "EITHER *A* OR *B*", or "EITHER *A* OR *B* OR BOTH." etc. Why is it unreasonable to use standard logic or standard probabilities to deal with the adequacy of such logical predications?

Problem of exceptions. If we take logical connectives to be defined by their standard truth tables (cf. Figure 2, later), a connective is traditionally considered to be false if there is inductively a single observation where this connective is false, or in a deductive context if there is a counter-example to an inference. However, particularly for conditionals this 'deterministic' interpretation has been challenged, in both philosophy and psychology (Adams 1986; Evans & Over, 2004; Oaksford & Chater, 2007; Oberauer, Weidenfeld, & Fischer, 2007; Over, Hadjichristidis, Evans, Handley, & Sloman, 2007). But also, other logical connectives may seem to require a probabilistic interpretation: predications about logical connectives based on empirical facts, such as "ravens are black AND they can fly," or "this pub is visited by people who are young AND male (or "by young, male customers"), appear to require tolerance of exceptions. Otherwise, most propositions in use would require abandoning due to being falsified (von Sydow, 2013). The ubiquity of exceptions is one of the reasons why it appears rational to replace (deterministic) logic as the adequacy criterion for logical predication by a high-probability criterion (Adams, 1986; Schurz, 2001, 2005; Foley, 2009). Yet still problems remain (von Sydow, 2011; von Sydow & Fiedler, 2012), of which only the most pressing is outlined:

Problem of inclusion: A fundamental problem for a direct use of extensional probabilities, P_E, as adequacy criteria for logical predication is that a more specific hypothesis could never be preferred over a more general one, even if it fits the data very well. For instance, the proposition "ravens are black AND they can fly" can never obtain a higher probability than that of "ravens are black OR they can fly or both" (an inclusive disjunction), since it generally holds that $P_E(B \wedge F) \leq P_E(B \vee F)$. Using extensional probability as adequacy criterion would thus never allow preferring the use of a conjunction over a more general inclusive disjunction. To carry this argument to its extreme, no specific hypothesis could be preferred over the general logical tautology (or the connective "verum", ⊤)—"Ravens are black or not, and they can fly or not (everything is possible)"—since its extensional probability of the tautology is $P_E(B \top F) = 1$. One thus could not prefer a specific connective over the general tautology (since $P_E(B \bigcirc F) \leq P_E(B \top F)$). Therefore, extensional probabilities as such, albeit useful in other contexts, are not informative measures to characterize the adequacy of a logical rule, given some data.

One important step toward solving the problem of inclusion is to use a subjective probability account together with strong sampling. For instance, Tenenbaum and Griffith (2001a) have argued, in another context, that with the additional assumption of strong sampling, Bayesian updating of subjective probabilities provides a natural penalty for more general hypotheses (cf. Navarro, Dry, & Lee, 2012). For weak sampling, the likelihood of the data given the hypothesis, $P(D|H)$, is 1, as long as all evidence falls into an acceptance region of H (otherwise it is 0). Although a Bayesian account as such (based on weak sampling) does not solve the problem of inclusion, things may differ for strong sampling. The latter assumes that the data is sampled from the consequence region, and thus the size of the consequence region s does matter (with $P(D|H) = 1 / s$ for an observation in the acceptance region and $P(D|H) = 0$ outside of it). This is also crucial for the principle of simplicity in perception, learning and high-level cognition (Chater & Vitanyi, 2003; cf. Lombrozo, 2007).[2] If we introduce strong sampling into the field of logic, it only leads to a *higher* likelihood and posterior probability for more *specific* hypothesis if the hypothesis with the smaller consequence region (e.g., the conjunction) has the *same* extension (no exceptions) as the larger one (e.g., the disjunction). This goes beyond a *strictly* extensional approach. But if there are observations specifically favouring the more general hypothesis (exceptions), such a subjective Bayesian account together with strong sampling alone does not solve the problem of inclusion (cf. Tenenbaum & Griffith, 2001b).

A pattern approach builds on strong sampling, but goes one step further, in proposing probabilistic measures for a kind of fit between a data-pattern and a logical explanation-pattern,

---

[2] Cf. the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Although they do not provide probabilities, they deal with the trade-off between goodness of fit and complexity of models (here, number of free parameters) (Burnham & Anderson, 2004).

while *allowing* for exceptions. As we will see, according to this approach a more specific hypothesis can be more probable than a more general hypothesis, even if there are exceptions.

Bayes Logic: A Sketch of the Model

Building on the outlined rationale, a possible computational-level model for such (intensional) pattern probabilities is now outlined. The model is an inductive logic, called 'Bayesian pattern logic' or 'Bayes logic', BL, for short (for further details see: von Sydow, 2011; cf. von Sydow 2013, 2014, 2015; von Sydow & Fiedler, 2012).

In contrast to standard extensional probabilities, P_E, BL formalises pattern probabilities, P_H, by providing a probability measure for the adequacy of alternative logical connectives. These probabilities are posterior probabilities formalising the belief in logical rules given the data (e.g., $P_H(A \wedge B \,|D)$). Here rules of increasing generality are not only nested sets – whose extension can never decrease – but also alternative hypotheses about explanatory logical patterns, where a more specific pattern can provide a more adequate explanation for a given data pattern than a more general pattern-hypothesis. For example, if one is concerned with the adequacy of the hypotheses that "animals of species *X* are aggressive AND curious" ($A \wedge B$), "they are EITHER aggressive OR curious" ($A >< B$), or "they are aggressive OR curious or both" ($A \vee B$), BL treats them as alternative pattern-hypotheses where, for instance, $P_H(A \wedge B|D) > P_H(A \vee B|D)$ is possible (Table 1, last row). The pattern-hypotheses will be formalised by probability tables, relating to logical truth tables and degrees of allowed exceptions. One can then calculate the likelihood of such pattern-hypotheses given the data, and use Bayes' theorem to calculate their posterior probabilities. We now proceed stepwise:

(1) *Input*. We are here concerned with a standard frequency input in a 2×2 contingency table with four possible events $A \wedge B$, $A \wedge \neg B$, $A \wedge \neg B$, and $\neg A \wedge \neg B$ (and the resulting cell frequencies $n_a + n_b + n_c + n_d = N$). To exclude memory effects, the observed instances will be explicitly presented here in a table (Fig. 1).

Insert Figure 1 about here

The model is based on the assumption that each event in the four cells of the table can occur independently, at least in principle, and it models a belief-update of believed (not ideal) cell probabilities based on observed joint frequencies.[3] However, using this input does not rule out original misrepresentations of frequencies, occurring by causes exogenous to the model. For instance, people may sometimes not have a full representation of observed frequencies and calculate them from estimated marginals only. In complex settings this seems plausible, and BL leaves room for other theories' modelling such subjective deviations from objective frequencies. BL simply takes *subjective* frequencies as input. These subjective joint frequencies may even be used to describe subjectively represented joint *probability* distributions, even if frequencies are not explicitly available (von Sydow, 2013; cf. Nilsson, 2008). In any case, here we concentrate on situations with transparent frequency input.

(2) *Ideal logical probability tables* (PTs) are logical-but-noisy patterns that may potentially explain the observed data. The PTs correspond to particular connectives in analogy to their standard truth-table definition (Figure 2, Column 1). Each PT as a whole has a (second-order)

[3] The belief update for the subjective *cell* probabilities uses a Dirichlet distribution for independent events, corresponding to the multinomial distribution in Step 3 of the belief-update of ideal pattern-hypotheses. This flexible distribution, however, does not resolve the problem of inclusion, but only of the use of pattern-hypotheses.

probability. A PT itself is a pattern-hypothesis built from four probabilities with values between [0, 1] and $p_a + p_b + p_c + p_d = 1$. These four cell probabilities are formalised as probabilistic analogues to the truth tables of dyadic logical connectives, $A \bigcirc_l B$. Figure 2, for instance, shows PTs for the conjunction '*A* AND *B*,' the dyadic affirmation *A*, '*A* (and *B* or non-*B*),' the dyadic affirmation *B*, '*B* (and *A* or non-*A*),' the exclusive disjunction 'either *A* or *B*,' and the inclusive disjunction, '*A* OR *B* or both.'

Insert Figure 2 about here

Assumption of idealization: Figure 2 illustrates that the cell probabilities of these PTs, unlike observed cell probabilities, are ideal explanatory hypotheses. These ideal PTs are assumed to have equal cell probabilities for cells that are logically true (*T*) according to a connective's truth-table definition (Figure 2, first column; cf. Johnson-Laird et al., 1999, in the context of logic; and Tenenbaum & Griffith, 2001a, in the context of Bayesian models). By contrast, standard logic makes no assumption about the ideal distribution of true cases. The shadings in the PTs of Figure 2 indicate assumed cell probabilities. For PTs with a 'deterministic' zero-tolerance for exceptions ($r = 0$, second column of Figure 2), the (ideal) probability of a cell that is logically true (*T*) is thus claimed to be t, obtained by dividing 1 by the number of true cells in the truth table of a given connective (strong sampling). This holds for the conjunction (one true cell) that $t = 1$, for the dyadic affirmation and the exclusive disjunction that $t = 1/2$, and for the inclusive disjunction that $t = 1/3$. Thus a 'deterministic' PT ($r = 0$) for an inclusive disjunction would be [.33, .33, .33, 0], and for a conjunction, [1, 0, 0, 0]. After observing *A*&*B* cases only, the likelihoods (cf. Step 3) of the conjunction could thus be higher than for the inclusive disjunction: ($P(D|PT)$ = .33 vs. 1). However, for $r = 0$, PTs are still falsified by single counter-examples.

Assumption of uncertainty: If one tolerates exceptions ($r > 0$), PTs represent connectives at a particular level of ideally distributed noise. Figure 2 illustrates PTs for some connectives at three noise-levels only (our predictions are based on modelling 11 equidistant levels, but 101 levels led to the same predictions). The PTs can be understood as a mixture of two distributions, and r (with $0 \leq r \leq 1$) specifies to what extent either the deterministic logical PTs or an ideal noise distribution (equiprobable cell probabilities, each with $p = .25$) determine the resulting PTs (von Sydow, 2011).[4] Table 2 shows numerical examples for the resulting PTs for an assumed intermediate level of $r = .5$.

Insert Table 2 about here

(3) *Likelihood*. For a PT (a hypothesis $H_{l,\, r,}$ combining a connective, l, with an assumed noise level r), the likelihood of the observed data D (the four frequencies n_a, n_b, n_c, n_d) is to be calculated by a multinomial distribution using the four probabilities constituting a PT.

$$P\left(D \middle| H_{l,r}\right) = \binom{N}{n_a n_b n_c n_d} {p_a}^{n_a} {p_b}^{n_b} {p_c}^{n_c} {p_d}^{n_d} \quad (1)$$

For an observed data pattern D, e.g. [20, 10, 10, 10] or [17, 5, 2, 2], and a particular assumed noise level (e.g., Table 2), the resulting likelihoods L of the conjunction are larger than both the likelihood of the affirmations and the likelihood of the inclusive disjunction (here $L(A \wedge B) > L(A) > L(A \vee B)$). This is driven by the likelihood function based on the PTs.

---

[4] Formally, the ideal cell probability of a logically false cell is $.25 * r$ (with a maximum of .25) and that of a true cell is $t - r\ (t - .25)$, with t as defined above.

(4) *Posterior*. The likelihood $P(D|\, H_{l,\, r})$ of the data D given an ideal PT (that is, $H_{l,r}$) can then be turned around, by Bayes' theorem, to obtain the pattern-based posterior probability of a PT given the data, $P(H_{l,\, r}\, |D)$, while here assuming a uniform prior for all PTs.

$$P(H_{l,r}|D) = \frac{P(D|H_{l,r})P(H_{l,r})}{P(D)} \quad (2)$$

(5) *Summation*. The posterior probabilities of the PTs of a connective l are summed up over all noise-levels r, resulting in pattern probabilities for dyadic connectives, $P_H(A \text{ o}_l\, B|D)$; that is:

$$P_H(H_l|D) = \sum_{\boldsymbol{r}} P(H_{l,r}|D) \quad (3)$$

(6) Additionally, to account for sensitivity to odds rather than to differences, the average log-odds of all pairs of compared hypotheses may be calculated to account for differences between small pattern probabilities (cf. Anderson, 1990; Griffiths & Tenenbaum, 2005).

A Bayesian Model Predicting a System of Logical Inclusion Fallacies

Bayesian pattern logic is a non-extensional account of Bayesian belief-update that uses the idea of logical patterns-hypotheses with different noise-levels to solve simultaneously the sketched problems of inclusion and exception. In this paper the pattern approach of CF is generalized to several other connectives, and the results are compared to several other approaches.

Since CFs have been a major anomaly for a Bayesian account of cognition, the computational level model of BL may contribute to the Bayesian renaissance in cognitive science, by which it has been inspired (e.g., Anderson, 1990; Chater & Oaksford, 2008; Griffiths & Tenenbaum, 2005; Goodman, Tenenbaum, Feldman, & Griffiths, 2008; Hahn & Oaksford, 2007; Hartmann & Meijs, 2012; Kemp & Tenenbaum, 2009; Kruschke, 2008; Oaksford & Chater, 1994, 2007; Tenenbaum & Griffith, 2001a,b). Some other Bayesian models have likewise intended to bridge the gap between Boole and Bayes, but they differ in technical detail and with regard to the tasks modelled (von Sydow, 2011). For instance, Oaksford & Chater's (2003) explanation of Wason's Selection Task (WST; with exception parameter) was a primary inspiration for BL, and was formulated in the context of logics. However, it did not address the problem of inclusion and does not suggest a probabilistic model of conjunctions.[5] Although it has justly been argued that Bayes' theorem alone cannot rationally justify violations of the conjunction rule (Gigerenzer, 1991, 1998; Fisk, 1996), BL provides an intensional pattern-based account of CFs and other IFs by building on strong sampling and addressing the problems of inclusion and exception simultaneously.

Thus, with reservations, BL is claimed to be a descriptive model as well. *If* people aim at judging the descriptive adequacy of alternative dyadic connectives using probabilities, judgements are claimed to be at least approximately coherent with BL. Although the increasing popularity of ideas such as 'Bayesian brains' (e.g., Doya, Ishii, Pouget, & Rao, 2007) lends credence to the suggestion that BL might serve as an algorithmic or implementation-level account, I would not exclude that a fast and frugal "pattern heuristic" could implement pattern probabilities (Gigerenzer, Todd, & the ABC Research Group, 1999; cf. General Discussion).

The pattern probabilities, P_H (here modelled without any free parameters), are formulated using standard probabilities, but they show 'emergent properties' that differ from a direct application of extensional probabilities. Most crucially, BL predicts occurrences of CFs and other inclusion

---

[5] Von Sydow (2006) proposed generalising the WST model, traditionally limited to conditionals, to other connectives, but that account differs from the present one as well.

fallacies (IFs). Assuming flat priors and, for instance, frequencies corresponding to an AND pattern (e.g., [20, 10, 10, 10], or [17, 5, 2, 2]), BL predicts double conjunction fallacies, $P_H(A \wedge B \mid D) > P_H(A \mid D)$, $P_H(A \wedge B \mid D) > P_H(A \mid D)$ (excluding several alternative theories of the CFs). If data is more likely produced by an ideal affirmation *A* (e.g.: [20, 20, 10, 10] or [13, 8, 2, 3]; cf. Experiments), the affirmation *A* should be judged the most probable hypothesis involving single CFs, $P_H(A \mid D) > P_H(A \wedge B \mid D) > P_H(B \mid D)$.

In the classic Linda conjunction 'fallacy' task, answer-patterns similar to the latter example were most frequent (Tversky & Kahneman, 1983). In the Linda task, participants first read a story about Linda may be taken to suggest that she was a feminist. Then participants ranked several statements by their probability: "Linda is an active feminist" (*A*); a 'bank teller" (*B*); and "a bank teller and active in the feminist movement" (*B* and *F*). The resulting belief pattern may correspond to the last mentioned frequency patterns if people have to imagine 'several Lindas.' Even if one assumes a low marginal probability for *B* (bank teller), with frequency estimations such as [10, 35, 5, 25], BL would predict the found dominant single CFs.

But the specific variant of BL only applies if people are concerned with *dyadic* connectives dealing with *dichotomous* events; and in other contexts other variants of BL may apply. First, BL's assumption of equiprobability in an affirmation models a *dyadic* affirmation *A* that involves a representation of *B* as well. There may well be contexts where an alternative monadic interpretation, without representation of a *B vs. non-B* subclass, is used instead (and one reviewer made a similar suggestion). I have elaborated elsewhere that not only the affirmation, but also the ordinary language 'AND' may often refer to a conjunction of monadic affirmations instead of dealing with a dyadic relationship (von Sydow, 2014). But in situations as in the present experiments, a dyadic interpretation seems most appropriate, since *B* clearly belongs to the context, we use dyadic formulations, and we compare a hypothesis with several other dyadic connectives. Second, there seem to be contexts in which investigated attributes are not represented dichotomously, but polytomously, involving several subclasses and perhaps implying another model variant (von Sydow, 2015; cf. General Discussion). Here, however, we focused on dichotomous event interpretations–for instance, by explicitly providing a 2×2 data input.

Some aspects of (dichotomous, dyadic) BL, such as sample-size and pattern-sensitivity, have been tested before (von Sydow, 2011); and in preliminary publications BL has been corroborated in sequential learning scenarios (von Sydow & Fiedler, 2012), even without providing explicit frequencies (von Sydow, 2013).

Here we focus on one central prediction fundamental to the general idea of logical pattern probabilities. If BL provides a roughly correct explanation for the class of pattern-based CFs, it should be possible to embed CFs in a system of logical inclusion 'fallacies' (IFs), even if based on frequencies in summary tables. BL predicts several other IFs, such as judging the conjunction $A \wedge \neg B$ to be more probable than the exclusive disjunction 'EITHER *A* OR *B*', $P_H(A \wedge \neg B) > P_H(A >< B)$. Likewise, the exclusive disjunction may become more probable than the inclusive disjunction, '*A* OR *B* or both' ($P_H(A >< B) > P_H(A \vee B)$).

Although aiming hereby to contribute to a more complete 'periodic table' of IFs, the experiments reported largely exclude conditionals. If conditionals always referred to a fully represented material implication, or to an equivalence relation, BL would be directly applicable. However, conditionals appear to be more directed (often with asymmetries to the logically equivalent counterpositive, $A \rightarrow B \Leftrightarrow \neg B \rightarrow \neg A$). Moreover, conditionals often appear *not* to be concerned with a full 2×2 probability table, often only expressing a high *P*(consequent |

antecedent) (Kao & Wassermann, 1993; Johnson-Laird & Byrne, 2002; Over & Evans, 2003; Evans & Over, 2004; Novick & Cheng, 2004; Griffith & Tenenbaum, 2005; Oberauer et al., 2007; Over et al., 2007; von Sydow, 2014b).

In addition to testing situations with several logical connectives and investigating whether people use something like pattern probabilities to determine the *most* probable hypothesis, it also should be investigated whether they use pattern probabilities when determining the *second-most* probable as well.

Conjunction Fallacy Debate

As background for our experiments, we briefly consider three important aspects of the intensive debate on the conjunction fallacy**.**

(1) It has been reasonably argued that the fallacy may be due to *misunderstandings of logical terms*. The pragmatics of conversation may lead participants to interpret the ordinary language phrase “is a bank teller” (*B*), if contrasted to “bank teller and feminist” ($B \wedge F$), as “bank teller and not feminist” ($B \wedge \neg F$) (Hilton, 1995; Mellers et al., 2001). Additionally, “bank teller and active feminist” may be interpreted, not as a logical conjunction, but as a sum or logical disjunction, ‘$B \vee F$’ (Hilton, 1995; Mellers et al., 2001; Hertwig, Benz & Krauss, 2008; von Sydow, 2014, 2015; but cf. Tentori & Crupi, 2012). Even though such interpretations plausibly play a role in exculpating participants from having committed some CFs, here it should be shown that there is a class of pattern-based inclusion fallacies independent of such interpretational considerations. We use specific formulations to ensure that people interpret ordinary language in the sense of ideal dyadic connectives: for the affirmation ‘*B,*’ we use “Linda is a bank teller, whether or not she is active in the feminist movement” (Messer & Griggs, 1993; Tversky & Kahneman, 1983);

and for the logical conjunction a formulation clearly referring to intersections: “bank tellers *and at the same time* active feminists” (cf. von Sydow, 2014; Messer & Griggs, 1993).

(2) The interpretation of *“probability”* has come under closer scrutiny as well. We will test some of the accounts that appear applicable to our frequency-based experiments with several alternative connectives. (a) The *inverse probability approach* claims that subjects confuse the target judgements of a conditional probability, $P(B \wedge F \mid L)$, with its inverse, $P(L \mid B \wedge F)$ (Wolford, 1991; Fisk, 1996; cf. Hertwig & Chase, 1998). (b) *Averaging accounts* of CFs posits that subjects use the average of the probability of the two conjuncts to arrive at judgements about the combined conjunction: $P(B \wedge F \mid L) = \text{mean}(P(B|L), P(F|L))$ (e.g., Gavanski & Roskos-Ewolsen, 1991; cf. Wedell & Moro, 2008; Jarvstad & Hahn, 2011; Jenny, Rieskamp, & Nilsson, 2014). But averaging-approaches, just as other rule-combination approaches (cf. Hertwig & Chase, 1998), reasonably apply “only when the probability of the conjunction cannot be assessed directly” (Nilsson, 2008, p. 473). In the present experiments this is not the case; thus here we only explore an alternative “averaging heuristic”, applicable to all investigated connectives: the probability of a dyadic connective is postulated to be the average of the observed probabilities of its true cells (e.g., $P_{AV}(F) := (P(B \wedge F) + P(\text{non-}B \wedge F)) / 2$). (c) Finally, confirmation accounts of the CF (e.g., Lagnado & Shanks, 2002; Sides et al., 2002; Crupi, Fitelson, & Tentori, 2008; cf. Schupbach, 2009; Tentori, Crupi, & Russo, 2013) formalize the idea that probability judgments may in fact often be judgements about an *increase* of probability (confirmation judgments). For instance, in the classical Linda tasks, the probability of a conjunction $B \wedge F$ is assumed to be answered falsely by confirmation judgments where $c(B \wedge F \mid L) > c(B \mid L)$, even if such an inequity cannot hold for standard probabilities. Here only two common measures of confirmation c will be modelled, a

difference and a ratio measure: $c_D(B \wedge F) = P(B \wedge F \mid L) - P(B \wedge F \mid \neg L)$ (e.g., Lagnado & Shanks, 2002); and $c_R(B \wedge F) := P(B \wedge F \mid L) / P(B \wedge F \mid \neg L)$ (e.g., Sides et al., 2002). Although this approach has not been applied to a general system of connectives before, we do so here.

(3) Other proposals relate to effects of *presentation formats*. Gigerenzer, Hertwig, and colleagues (Gigerenzer, 1996, 1998; Hertwig & Gigerenzer, 1999; Hertwig & Chase, 1998) have argued that CF-effects may concern single-event probabilities only. This would involve a violation of subjective theories of probability but not of the traditionally major view of probability—the frequentist conception. In fact, CFs were shown to be mitigated by tasks referring to natural-frequency formats that report the final tally of a natural sampling process (Tversky & Kahneman, 1983; Kleiter, 1994; Fiedler, 1988; Cosmides & Tooby, 1996; Gigerenzer, 1996, 1998, 2000; Kahneman & Frederick, 2002, 2005). Furthermore, there is evidence for perhaps alternative facilitation effects due to transparent set inclusions – for instance, caused by removal of filler items (Tversky & Kahneman, 1983; Mellers et al., 2001; Sloman & Over, 2003; Sloman et al., 2003; Neace et al., 2008). However, there have been striking examples of CFs even in situations with both frequency information and clear set inclusions (Lagnado & Shanks, 2002; Sides et al., 2002; Tentori, Bonini, & Osherson, 2004; Wedell & Moro, 2008; von Sydow, 2011). Although I think that frequency information and clear set inclusions may often facilitate adequate extensional probability judgments, they seem no sufficient conditions for extensional thinking, even jointly. The aim here is to elicit IFs in quite simple, fully transparent tasks, with transparent, full frequency information about two features in contingency tables, with no filler items, no memory load, and no qualitative story. This excludes a number of explanations of CFs, independent of their validity in other contexts (e.g., Betsch & Fiedler, 1999; Hartmann & Meijs, 2012; Hintikka, 2004; Lagnado & Shanks, 2002; Sloman et

al., 2003; Wedell & Moro, 2008; Jarvstad & Hahn, 2011). In the present paper it is argued that if one is concerned with the adequacy of alternative logical explanatory patterns, extensional relative frequencies, $P_E(B \wedge F|L)$, do not provide the requested target measure. Instead, pattern probabilities–$P_H(B \wedge F|L)$—should apply, predicting a system of inclusion fallacies even if people are confronted with frequencies and clear set inclusion.

Testing Systematic Logical Inclusion Fallacies

The present studies aim to provide an existential proof that there is a class of logical inclusion 'fallacies' linked to logical patterns, even under strict conditions using several extensional cues. We also aim to rule out that the other CF accounts sketched in the introduction, such as inverse probability or two measures of confirmation, can explain our findings, without arguing that there are no other domains of application for such approaches.

Apart from an intensive debate on conjunction fallacies, there have been relatively few studies on inclusive disjunctions (e.g., Tversky & Kahneman, 1983; Biela, 1986; Carlson & Yates, 1989; Bar-Hillel & Neter, 1993; Reeves & Lockhart, 1993; Costello, 2005; Costello & Mathison, 2014). Most of the studies used a *qualitative story* to vary the representativeness of the constituents, and only some varied the probability of the composing events (Carlson & Yates, 1989; Costello, 2005; cf. Thüring & Jungermann, 1990; Bar-Hillel & Neter, 1993). Here we provide frequency information and explicit feedback on all four cells of the contingency table, and involve not only inclusive disjunctions but also exclusive disjunctions in the hypotheses space.

Experiment 1 tests several logical inclusion fallacies, by giving information about observed frequencies in contingency tables rather than by verbally encoded scenarios. Participants should select the hypothesis with the highest probability from among several logical hypotheses so that, for instance, selecting a single AND hypothesis involves several IFs, predicted by BL, such as $P(B \wedge F) > P(B) > P(B \vee F)$ or $P(B \wedge \neg F) > P(B >\!\!-\!\!< F) > P(B \vee \neg F)$ (von Sydow, 2009; cf. von Sydow & Fiedler, 2012).

Experiment 2 continues to test the system of IFs predicted by BL. Here, however, the frequencies observed are less clear-cut from the perspective of BL, resulting in *two* hypotheses with high pattern probabilities. First, it is investigated whether participants select the most probable hypothesis according to BL in such situations. Second, participants judge which is the second-most probable hypothesis, investigating whether participants continue to use pattern probabilities for these second judgements.

In contrast to most standard CF tasks, experiments presented here did not use stories about Linda, but instead, transparent frequency information about a class of people—e.g., pupils of a Linda school (cf. Lagnado & Shanks, 2002, for more complex frequency-based learning tasks). Thus memory effects are ruled out as major mediators or moderators, due to providing full frequency information about the co-occurrence of the two attributes in a contingency table.

Experiments 1a and 1b: Simultaneous Logical Inclusion Fallacies

Experiments 1a and 1b are similar versions of an experiment to elicit pattern probabilities in a more or less explicit way. Both studies followed the same procedure, with almost identical material.

Participants

In Experiments 1a and 1b, 64 students from the University of Göttingen (32 per experiment) volunteered to participate.

Procedure

In both experiments, the participants received booklets including general instructions, four probability-judgement tasks, and a final questionnaire concerning demographic data and comments. The four probability judgement tasks concerned graduates of Maria, Linda, Magda, and Xenia schools, each with frequency information in a contingency table about a sample of graduates. Table 3 reported the frequencies for each school. In Experiments 1a and 1b we use data patterns with exceptions, that from the perspective of BL should nonetheless elicit different logical hypotheses in a relatively clear-cut way. Participants were randomly assigned to different presentation-orders of the schools (Latin square, Table 3), affecting the predictions of several of the modelled alternative approaches.

Materials

Each probability-judgement task concerned a different single-sex school (e.g., the Linda school). In order to avoid an extensional concern with subsets, an interest in the overall evaluation of the situation at the different schools was suggested to the participants. There was no supplemental information about these schools—merely frequency information about the co-occurrence of the attributes in question, based on a sample of graduates from each school. Information included the number of graduates who became ‘bank tellers and active feminists,’ ‘bank tellers and no active feminists,’ ‘no bank tellers and active feminists,’ and ‘no feminists and no active bank tellers.’ The occurrences in each cell of the contingency table were represented by pictorial symbols (‘smileys’; cf. Figure 1).

Insert Figure 1 about here

Both the cells and the marginals of the contingency table were labelled, to prevent confusion concerning set inclusions (Fig. 1; cf. Appendix A; Kahneman & Frederick, 2002; Kahneman & Frederick, 2005; Sloman et al., 2003; Johnson-Laird et al., 1999; Girotto & Gonzales, 2001).

Participants were then to judge intuitively which hypothesis was most probable, according to the situation described. One, for example, was that the Linda school graduates would generally become 'bank tellers AND at the same time active feminists' (*A* and *B*); another, 'bank tellers, whether they are feminists or not' (*A*). As mentioned earlier, the formulations were suggested in the literature to prevent logical misunderstanding. No 'filler hypotheses' were used. Participants chose the most probable from ten logical hypotheses about a given graduate school, referring to A, B, non-A, non-B, $A \wedge B$, $A \wedge$non-B, $A \vee B$, $A \vee$non-B, $A >< B$, and $A \rightarrow B$ (cf. Appendix A, Figure 1). We selected this relatively large but limited number of connectives since the propositions should refer to several truth tables often used in ordinary language, while not making the search among too many options too demanding.

Predictions

Predictions for Experiments 1a and 1b for BL (Figures 3A, D) and several alternative accounts of the CF—including extensional probability, averaging heuristic, inverse probability, and the difference and ratio confirmation measure (Figure 3 D)—were derived, based on the observed frequencies shown and according to the formulae mentioned in the Introduction (for Experiment 2, see Figures 4D, E). BL, extensional probability, and averaging heuristic make predictions, independent of order, as to which selected hypotheses should dominate for each school.

Some additional comments are needed with regard to inverse probability and confirmation which are concerned with measures relative to previous experiences. Inverse probability is not concerned with the probability of a connective given a school S, $P_E(A \circ B|S)$, but rather with the probability that someone belonging to the described set attends this particular school: $P_{INV}(A \circ B) = P_E(S|A \circ B) = P_E(S|A \circ B) / \Sigma_i P_E(S_i|A \circ B)$). Likewise, confirmation (in fact not ranging between 0 and 1) is here modelled as either $c_D(A \circ B|S) = P_E(A \circ B|S) - P_E(A \circ B|\neg S)$ or $c_R(A \circ B|S) = P_E(A \circ B|S) / P_E(A \circ B|\neg S)$) and is essentially comparative. In their explanations of traditional Linda tasks, both the inverse probability and the confirmation accounts normally assume that, prior to any evidence, P_E(bank teller and feminist) << P_E(bank teller) ($P_E(B \wedge F) < P_E(B \wedge \neg F)$). In this case, it holds for inverse probability that $P(\text{Linda}|B \wedge F) > P(\text{Linda}|B)$, and it holds for confirmation that $c(B \circ F|S) > c(B|S)$. Applied here to all schools presented first, these accounts would thus have to make the prediction that *B*-AND-*F* hypotheses should dominate the *B* hypotheses even for the Maria schools (with observed $f(B \wedge F) \approx f(B \wedge \neg F)$), where BL predicts no dominant *B* selections. If, alternatively, one assumed *equal* expectations for all cells in the table, inverse probability and confirmation theory would still have to predict many double CFs in these first Maria schools of Experiments 1 and 2. A further option is to derive predictions based on the data alone (only applicable after Position 1). Due to the four presentation orders, this would imply a prediction of strong serial-order effects (see Figure 3D; cf. Figures 4D, E).

Difference between Experiment 1a and Experiment 1b

After running Experiment 1a, Experiment 1b aimed to replicate the findings, using stricter conditions. These simple experiments differed in two respects only. (1) Experiment 1a tested whether IFs occur under the qualitative settings used (frequency information, etc.) when the use of pattern probabilities is clearly suggested to the reader. By contrast, Experiment 1b tested whether pattern probabilities occur in situations in which no direct suggestion for a pattern interpretation was made (Appendix A). To elicit pattern probabilities clearly (despite several extensional cues), Experiment 1a included the following formulations: “You are not interested in particular subgroups, but in different hypotheses (*H*) investigating the overall situation at these schools”; and “the hypotheses are concerned with the overall situation at this particular school.” Experiment1b only suggested such a reading implicitly. In both experiments, participants had to choose the most probable hypothesis, given the situation.

(2) Additionally, the formulation of the logical disjunction was changed after obtaining the results of Experiment 1a (which were positive, but with a relatively low number of OR-selections). Experiment 1a used: “The graduates […] become *bank tellers OR also active feminists* or also both.” In Experiment 1b, the formulation stressed set inclusions even more clearly. For instance, H7 read: “The graduates […] become feminist *bank tellers, OR ALSO non-feminist bank tellers, OR ALSO feminist non-bank tellers*.”

Results for Studies 1a and 1b

The main results of Experiments 1a and 1b are shown in Figure 3 B/C, here conflated over presentation order. They show the proportions of participants selecting different hypotheses in the four scenarios in both studies. The comparison to the pattern probabilities predicted by BL (Figure 3A) shows a high fit with the results, even visually (but also note the different scaling of Figures 3A, 3B and 3C, and that not all participants follow the predictions of BL).

In *Experiment 1a* the mode of the participants’ selections (the hypothesis regarded as most probable) corresponded in all schools to the answer predicted by BL (Figure 3B). For each school, the proportion of predicted answers clearly surpassed the chance-level (1/10: one hypothesis from ten alternatives; Maria: $\chi^2(1, n = 31) = 156.6$, $p < .0001$; Linda: $\chi^2(1, n = 31) = 141.9$, $p < .0001$; Magda: $\chi^2(1, n = 31) = 35.1$, $p < .0001$; Xenia: $\chi^2(1, n = 32) = 122.7$, $p < .0001$).

No other model (averaging, inverse probability or confirmation) predicted this pattern of results (Figure 3D).

Insert Figure 3 about here

Considering more closely the results for extensional probability, the predominant AND-selections in the Linda schools involve five IFs, since extensionally $P(A \wedge B)$ was estimated to be more probable than $P(A)$, $P(B)$, $P(A \vee B)$, $P(A \vee \neg B)$, and $P(A \rightarrow B)$. In the critical schools with different predictions for these two measures, 72% of the selections conformed with BL, whereas only 7% were predicted by extensional probability (with 21% reflecting other answers). In the Magda condition, in which BL as well as an extensional account predicted OR-selections, we obtained a relatively low confirmative proportion (41%) in Study 1a. This may be due to either a less clear BL-prediction for this pattern (Figure 3A) or an unclear "or both" formulation in H7 and H8 (cf. Figure 3D).

The overall results for *Experiment 1b* are shown in Figure 3C. Here, the use of pattern probabilities was suggested only implicitly, and a clearer OR-formulation was used. Overall, the proportion of predicted answers remained dominant. In the schools where BL differed from extensional probability in its predictions, 73% of the answers corresponded to the choice predicted by BL; 8% to that corresponding to extensional probability; and 19% of answers differed from both. Again, in all schools, the predicted BL-selections emerged at clearly above chance-level (Maria: $\chi^2(1, n = 32) = 165.0, p < .0001$; Linda: $\chi^2(1, n = 32) = 213.6, p < .0001$; Magda: $\chi^2(1, n = 32) = 66.1, p < .0001$; Xenia: $\chi^2(1, n = 32) = 76.1, p < .0001$).

Furthermore, the results of *both studies* differed markedly from the predictions of all modelled alternative theories (Figure 3D). Clearly, neither extensional probability nor the averaging heuristic could have predicted the results. Inverse probability and confirmation make comparisons between different schools, S_i, and cannot explain the results. As discussed in the prediction section, for the first presentation of a school one may employ either assumptions used to model traditional Linda tasks (based on a low $P(A \wedge B|\text{non-}S_i)$) or equal expectations for all four cell probabilities. In either case, one would expect dominant AND-selections, not only in Linda tasks but also in the Maria conditions. Empirically this was not found.

Additionally, these theories were modelled based on the data only, beginning with the second school. The resulting dominant predictions are presented in Figure 3D. To calculate the model-fits for the different presentation positions, the data from Studies 1a and 1b were conflated, resulting in a higher number of cases (n = 16 for each school at a particular position). Confirmation and inverse probability required a separate calculation for each presentation position (cf. Table 3). For instance, modelling the Linda school in the second position involved comparing it to the Maria school (Order W in the Latin square, Table 3).

Insert Table 3 about here

Table 4 shows the correlations between the predictions of the models (regarding which hypothesis *should* be selected) and the relative frequency of the hypothesis *actually* selected. In these correlations we considered the schools at each position (using conflated data from Studies 1a and 1b). The predictions of the models may be modelled either deterministically (hypothesis with a modal prediction is coded with 1, and the others with 0) or probabilistically (assuming hypothesis selection as proportional to the model's output). For BL, extensional probability, and averaging heuristic, one may calculate *global correlations*, conflating over positions and correlating the model predictions with the empirical findings from 40 cells only (10 hypotheses × 4 schools). However, inverse probability and the two confirmation measures, as mentioned, yielded data-based predictions from Position 2 onward only, and required considering presentation-order. Thus, here, correlations between each model and the relative frequencies found were calculated on this *detailed* level over 120 distinct cells (10 hypotheses × 4 schools × 3 positions). To obtain a comparable measure, this was also done for all models (‘detailed model’).

Insert Table 4 about here

The results in Table 4 show positive correlations, not only for BL but also for the other models, with particularly good results for averaging heuristic and confirmation_D. The positive correlations may have partly been due to an overlap of the predictions of the alternative models with those of BL. Yet participants seemed to use the dominant BL strategy only slightly more consistently than other strategies. We calculated the average correlations between the predicted (coded as 1) or unpredicted (0) answers in the four scenarios for each of the models: $r_{BL} = .21$, $r_{EX} = .03$, $r_{AV} = .01$, $r_{INV} = –.01$, $r_{COD} = .07$, $r_{COR} = .12$. However, the predictions of BL by any measure (probabilistic or deterministic model, without or with log odds, detailed or global fit-ting), corresponded to the data at a substantially higher rate than extensional probability, averaging heuristic, inverse probability, and confirmation (Table 4). Testing for the difference between correlations yielded significant results even for the most problematic comparison, $p < .001$.

Study 2: BL and Successive Rankings

This study aimed to show that participants use BL not only in situations where a particular hypothesis is predicted, but also for less clear-cut data, where a second logical pattern has a relatively high pattern probability as well. Furthermore, we asked participants to select not only the most probable but also the second-most probable hypothesis. We investigated whether participants continue to use pattern probabilities. Alternatively, such a successive assessment of hypotheses may trigger reflection and thus elicit an extensional norm, since it is presumably the only learned norm of probability judgment available for deliberate use (Stanovich & West, 2008; cf. De Neys, 2006; Evans, 2003; Osman, 2004).

Additionally, we again modelled the previously outlined predictions of potential alternative accounts: extensional probability, averaging heuristic, inverse probability, and confirmation (in difference- and ratio-measures).

Participants

Participants who volunteered were once again students from the University of Göttingen (*N* = 32), who as recompense were given a small gratuity and the chance to enter a lottery.

Procedure

As in Study 1, participants received a booklet with four probability judgement tasks concerning four different schools (Linda, etc.). They were randomly assigned to four presentation-orders (Latin square; cf. Study 1).

Materials

The instructions largely corresponded to those in Study 1. Participants again had to make intuitive probability judgements regarding graduates of four different high schools, based on data provided in contingency matrices rather than on stories or on complex learning procedures. For each high school, participants observed a sample of graduates (Figure 5, Panel A). Interest in an overall evaluation of the situation was encouraged. Participants tested only five hypotheses regarding the graduates, relating the attributes ‘bank teller’ (*B*) and ‘feminist’ (*F*) in different ways.

We used formulations analogous to those in Study 1b, linked to the following logical connectives: *B*; *F*; $B \wedge F$; *B*><*F*; $B \vee F$. Participants marked hypotheses as most and second-most probable for each situation.

Results

Panels E and F of Figure 4 show the hypotheses empirically ranked as most and second-most probable. The letters in the bar charts indicate the selections predicted to be dominant by different accounts. Figure 4, Panels B, C, D, shows the quantitative predictions for pattern probabilities (without a log-odd measure), extensional probability, and the averaging heuristic. It shows only these measures, since the predictions of inverse probability and confirmation, as in Study 1, depend on presentation order.

Insert Figure 4 about here

For all high schools, the most frequent selections of the most and second-most probable hypotheses corresponded to the predictions of BL. Again, selections in line with BL often involve several combined logical IFs. Nonetheless, the proportions of predicted answers in all four schools were above chance level (1 out of 5 hypotheses) for the predicted first selection (Linda: $\chi^2(1, n = 32) = 31$, $p < .0001$; Maria: $\chi^2(1, n = 32) = 36.1$, $p < .0001$; Magda: $\chi^2(1, n = 32) = 14.4$, $p < .001$; Xenia: $\chi^2(1, n = 32) = 33.3$, $p < .0001$) and for the second selection (Linda: $\chi^2(1, n = 32) = 26.3$, $p < .0001$; Maria: $\chi^2(1, n = 32) = 14.4$, $p < .001$; Magda: $\chi^2(1, n = 32) = 18$, $p < .0001$; Xenia: $\chi^2(1, n = 32) = 14.4$, $p < .001$). On average, 61% of the answers concerning the most probable hypothesis (in the critical Linda, Maria, or Xenia schools) were in line with BL, and only 18% with extensional reasoning (21% involved other answers). For the second-ranked group, 47% of answers (in critical schools) accorded with BL and only 20% with the extensional prediction (with the remaining 33% for other answers). Here, the averaging heuristic, which obtained good results in Experiments 1a and 1b, was clearly at odds with the results of two of the four schools (cf. Figure 4D).

As in Studies 1a and 1b, inverse probability and confirmation predict presentation-order effects. Again, these theories were modelled based only on the data shown. For instance, the prediction for presenting the Maria school in Position 3 is based on the frequencies for Positions 1 and 2. Furthermore, due to the four task-orders (cf. Table 3), the Maria school in the third position is always preceded by a Linda and a Magda school.

Table 5 shows for each model the correlation between their predictions and the actual relative frequencies of selections over schools. In order to provide separate fits for the hypotheses selected as most and second-most probable, we did not use rank-correlations, but rather calculated model-fits for the deterministic predictions for these selections separately. Deterministic predictions assume perfect sensitivity of the participants (coding predicted values by 1, the others by 0; cf. Study 1). For the hypotheses selected as *most* probable we again also calculated the models' *probabilistic* predictions and correlated them with the results, assuming selections proportionate to the probability predicted by the models. As in Studies 1a and 1b, we calculated not only the global correlations (aggregating over presentation-orders) but also the detailed correlations (considering presentation-order but not using the data for the first position).

Table 5 shows that all models—particularly confirmation$_D$—were positively correlated to the actual hypothesis selections, but again BL (either with or without the last model-step of calculating log-posterior odds) clearly achieved the best model-fit. The correlations for BL (in the detailed analysis) were significantly higher than those for the second-best models: For the first hypothesis selection, BL (without posterior odds) vs. confirmation$_D$ yielded $p < .01$ for the probabilistic model and $p < .001$ for the deterministic model. For the hypothesis selected to be second-most probable, the comparison between BL and the second-best model (here, extensional probability) likewise became significant, with $p < .01$. In conclusion, BL correlated reliably higher with the data than any other model tested.

Insert Table 5 about here

A high fit between predictions and data does not exclude that other strategies have been used consistently, at least locally. Local consistency seems more relevant than global consistency, since in conflict with traditional definitions of probability participants may change their strategy partway through an experiment. Experiment 2, in contrast to Experiment 1, allows accessing the *local* consistency within a task between first and second probability judgements for each scenario. The average correlation over the four tasks between a model's predicted (1) and unpredicted (0) selections in the two phases of a task suggests that some participants apply BL, but also other strategies, in a locally reasonably consistent way (r_{BL} = .60, r_{EX} = .40, r_{AV} = .-.33, r_{INV} = .35, r_{CD} = .17, r_{CR} = .49). In any case, the main results from Table 5 indicate that BL is favoured over other strategies in Experiment 2 as well.

General Discussion

The results obtained provide support for systematic occurrences of several logical inclusion fallacies coherent with pattern probabilities modelled here by BL (von Sydow, 2011).

In Experiments 1a and 1b, participants selected the most probable logical hypothesis from ten possibilities. The results provided evidence for a large number of different IFs. For instance, any selection of an AND-hypothesis involved four simultaneous IFs. Nonetheless, the vast majority of answers were coherent with the predictions of BL. In the experiment reported, conditions were used that involved a number of extensional cues, such as transparent tasks, explicit frequency information in contingency tables, no distractor items and no varied qualitative story. We obtained the predicted results in Experiment 1a, where these cues were counterbalanced by explicitly emphasizing an interest in the evaluation of an overall situation; but also in Experiment 1b, where such an interest was merely suggested.

Experiment 2 investigated shown frequencies for which the predictions of BL are less clear. BL here predicts two hypotheses to be quite probable, and it was investigated whether participants chose the most probable one. Moreover, it was tested whether participants acted according to BL if they also had to rank the second-most probable hypothesis. The results of Experiment 2 corroborated the predictions regarding both first- and second-ranked hypotheses.

In a further experiment not reported here (space precluding an exhaustive treatment), the same materials and frequencies were used, but we requested ranking all five logical hypotheses. Pattern probabilities still provided a highly significant influence, but it diminished significantly, whereas the number of extensional answers increased. This result supports the idea that considering several ranks (particularly if BL does not provide clear predictions for all rankings) may lead participants to use deliberately the known probability interpretation of relative frequencies instead of following their original intuition (cf. Stanovich & West, 2008).

In sum, the results corroborate a systematic occurrence of several logical inclusion fallacies and a combination of inclusion fallacies in two experiments asking for the most probable or the two most probable hypotheses, despite using transparent tasks, natural frequencies and clear set inclusions.

Quantitative and Qualitative Theories of CFs

The results cannot be explained easily by any other theory of the CF, including several qualitative and quantitative theories.

Quantitative Accounts

The results are not only at odds with standard extensional readings of probability, but they provide tests of some of the most promising alternative quantitative theories suggested in the CF debate. Here some of these theories were modelled with regard to a larger set of dyadic logical connectives (Figures 3, 4). These theories, plus a selection of influential further theories, will be discussed briefly. Many theories of the CF have been proposed as universal theories of CFs. The present results challenge such interpretations. However, some of these theories have plausibly distinct domains of application, even if they cannot account for the pattern-based IFs documented here.

(1) The *theory of inverse probability* (Wolford, 1991; Fisk, 1996; Fisk & Slattery, 2005; cf. Hertwig & Chase, 1998) posits that CFs occur due to confusion of the conditional probability, $P(B \wedge F \mid L)$, with its inverse, $P(L \mid B \wedge F)$. This theory can theoretically be extended to other logical connectives but was actually incoherent with several of the results presented.

(2) *Averaging, weighted averaging, signed summation,* and further combination rules have been used to explain traditional CFs (Gavanski & Roskos-Ewolsen, 1991; Carlson & Yates, 1989; Fantino, Stolarz-Fantino, & Wright, 1997; cf. Thüring & Jungermann, 1990) and have subsequently received some support (Nilsson, 2008; Wedell & Moro, 2008; Jenny et al., 2014; cf. Jarvstad & Hahn, 2011). Yet these accounts were not designed for use with all logical hypotheses tested here. Even more important, these theories were proposed for situations in which

the co-occurrence of joint events is less transparent than in the experiments reported here. Therefore a different ‘averaging heuristic’ was explored here, which – despite obtaining the second-best result in Experiment 1 –could not account for the overall systematics of the findings.

(3) Relatively recently, an innovative *quantum probability (QP) approach* for inclusion fallacies (at least for conjunctions and *inclusive* disjunctions) and other judgemental biases has been developed (e.g., Busemeyer, Pothos, Franco, & Trueblood, 2011). This approach applies mathematics of quantum physics to these biases but has not been extended to exclusive disjunctions as investigated here. The approach predicts CFs only when people do not easily represent a conjunction because two conjuncts are deemed “incompatible”. Thus QP seems applicable only if people have no detailed memory or accessible information about the conjunction of events, unlike the present experiments. Nonetheless we obtained systematic IFs. Thus if one applied QP here, it would not explain the results, since it has been postulated, as a “strong prediction of the quantum model,” that “only a single conjunction error can occur“—never a high portion of double CFs (Busemeyer et al., 2011, p. 202). This contrasts with the high proportions of double CFs found here.

(4) Costello and his colleagues (Costello & Mathison, 2014; cf. Costello, 2009) have recently elaborated a *probability-with-noisy-recall account* (PN-model), advocating that CFs and IFs may be based on standard probability estimates influenced only by random error in recalling events or means. PN seems similar to the earlier proposal of BL (von Sydow, 2007, 2009, 2011) in expecting IFs when there are not many exceptions to a subset-hypothesis. However, if the expected probabilities become similar (e.g., $P_E(A) \approx P_E(A \wedge B)$), the number of fallacies should increase due to *error-variation* around given means (e.g., resulting in $P_E(A) < P_E(A \wedge B)$ due to chance fluctuations). In contrast to BL, the maximum of predicted CFs thus appears limited to

50% at the utmost; in fact, "between 20% to 40%" CF rates would normally have been found after "controlling for various extraneous factors" (Costello & Mathison, 2014, p. 362). A recent study with a low 15-20% fallacy rate indeed provides substantial evidence for this model, particularly as it showed that $P(A) + P(B) – P(A \wedge B) – P(A \vee B)$ was symmetrically distributed around a mean of zero (Costello & Mathison, 2014). In my interpretation the repeated rating tasks employed triggered an extensional strategy (Hertwig & Chase, 1998; Sloman et al., 2003; cf. our briefly mentioned Experiment 3); and results indeed showed noise-based CFs. It seems unlikely, however, that this approach can explain the present findings, since our tasks reduce the plausible role of memory-error to a minimum; and it should be easy to determine the highest extensional probability (e.g., $P_E(A \wedge B) < P_E(A \vee B)$). The PN-model, moreover, does not account for the high rate of IFs, the systematic nature of our findings, or the results for successive ratings (Experiment 2). Additionally, other experiments have shown that people often committed CFs even when providing adequate frequency estimates themselves (von Sydow, 2013). Overall, although the PN-model does provide a plausible explanation of one class of CFs, it cannot convincingly explain the pattern-based CFs, here or elsewhere.

(5) *Confirmation* accounts of the CF (Sides et al., 2002; Lagnado & Shanks, 2002; cf. Crupi et al., 2008; Tentori & Crupi, 2009; Tentori et al., 2013) advocate that people substitute a measure for probability by one for the *increase* of probability. Although this is in my view plausible as *one* cause of CFs, it is not the only one and lacks a generally accepted applicability criterion when such substitutions are predicted (since there are clearly cases in which people use extensional probabilities; Beach & Peterson, 1966; Costello & Mathison, 2014).

Tasks such as ours, based on pre-classified frequencies of events in a 2×2 matrix, dissociate potential confirmation effects from effects potentially occurring *before* one arrives at such a dichotomous matrix. In contrast, if dichotomous events are not pre-classified and transparent, there may well be effects related to scale construction, dichotomisation of continuous events relative to a population mean or range, effects of anchoring and adjustment, and memory effects. These may all be indirect causes of CFs, and they need to be distinguished from confirmation. Nonetheless, confirmation (sometimes formalized as difference) has been supported, even using pre-categorised events (Lagnado & Shanks, 2002).

The concept of confirmation, although previously mostly limited to CFs, can naturally be extended to all logical connectives investigated here. However, the present, simple, transparent studies seem to favour a pattern approach over an (extensional) confirmation approach (cf. von Sydow, 2011).

One might object that the models of confirmation tested here include only two prominent formalizations—the subtractive one (Lagnado & Shanks, 2002) and the ratio-formulation (Sides et al., 2002; cf. Mahler, 2004). More recently, several other measures have been introduced into the psychological debate (Tentori et al., 2004, 2007; Crupi et al., 2008; cf. Maher, 2004; Nelson, 2005), and it is worth exploring them in their own right. They also substantially extend the flexibility and, perhaps, underspecification of a confirmation approach of CFs. Hence, it seems advantageous that Tentori, et al. (2013) have investigated sufficient but not necessary conditions where all confirmation accounts predict CFs (Crupi, Fitelson, & Tentori, 2008, 2009; critically, Schupbach, 2009).[6] These conditions, however, cannot directly be generalised to other connectives. The investigated measure could not explain the data as clearly as BL. Moreover, it is unlikely that even any other measure would have produced the effects obtained. Otherwise we would have obtained strong serial-order effects and less clear results. Moreover, in previous tests all measures turned out to be highly correlated (Tentori et al., 2007). Thus it seems fair to shift the burden of proof (of whether another measure of confirmation may explain the results) to those advocating such an explanation. For the time being, the current results strongly suggest the existence of a pattern-based explanation of IFs. Furthermore, von Sydow and Fiedler (2012) have shown in sequential-learning tasks that double CFs remained dominant even after *dis*confirmatory evidence (cf. von Sydow, 2015).

Recently, Tentori et al. (2013) provided substantial evidence in favour of confirmation-based CFs and against several other accounts of CFs. They assessed probability and confirmation judgements independently of CF tasks (we consider Experiments 1 and 2). Their results showed that participants most frequently assigned a higher probability to confirmed conjunctions rather than to those with an extensional probability above 50% (conditional on another event). Although these findings indeed rule out many CF accounts, they do not need to rule out a pattern account (pattern probabilities favouring the selections of cells that have a *relatively high* extensional probability). BL therefore could predict the selection of hypotheses whose extensional probability

---

[6] Apart from more specific conditions, they still advocate a confirmation approach, predicting CFs if and only if the confirmation $c(B \& F|e)$ is larger than $c(B|e)$ (or than $c(F|e)$. It is not clear to me whether a newly proposed condition, $c(B, F|e) > 0$, actually involves a deviation from this basic position (Tentori et al., 2013, p. 249).

is *below* .5. For instance, von Sydow (2011) corroborated a substantial number of double CFs even for data patterns such as [20, 10, 10, 10] (with an extensional probability, $P_E(A \wedge B\text{-cell}) = .4$). Moreover, Tentori et al.'s (2013) story-based tasks, even though concerned with dichotomized job categories (e.g., interpreter vs. no interpreter, Experiment 1), may elicit a more complex polytomous representation of several possible jobs – say 5 jobs [a, b, c, d, e] – with corresponding belief or frequency patterns, such as [16, 5, 3, 5, 4]. In such a situation, a version of polytomous BL (von Sydow, 2015) would predict that people's pattern probability for $P_H(X \text{ are } a)$ may be higher even within this single dimension than for $P_H(\text{X are non-}a) = P_H(\text{X are } b, c, d \text{ or } e)$, although $P_E(\text{a}) > P_E(\text{non-}a)$. We obtained first results with explicit polytomous representations that were incoherent with a confirmation approach but coherent with BL (von Sydow, 2015). Therefore, although the results of Tentori et al. (2013) exclude several other theories, they do not need to exclude a pattern account.[7] Despite this possible interpretation, and even after disconfirmation here, a confirmation approach (or a pattern-confirmation approach) arguably remains a plausibly valid account of another class of CFs.

---

[7] A further explanatory factor for the findings of Tentori et al. (2013, Experiment 2) may be relevant. After a Linda story, confirmation predicts $P(\text{bankteller} \wedge \text{feminist}) > P(\text{bank teller} \wedge \text{owns dark shoes})$, since the Linda story confirms the former and not the latter conjunction. However, one may have to differentiate relevance from confirmation: Imagine keeping the Linda story without changing the confirmation relations but changing the relevance of "black shoes", e.g., by inviting Linda to a party where one must wear dark shoes. In such cases, confirmation may invariably continue to predict *no* selection of "feminist & dark shoes" over "feminist & bank teller" and even over "feminist"; but, I would expect such selections would become more common.

(6) Finally, representativeness (Tversky & Kahneman, 1983) in its clarified formulation (Kahneman & Frederick, 2002; Kahneman & Frederick, 2005), combining the construction of a specific prototype and a similarity judgement, does not explain the results reported (cf. Nilsson et al., 2008). (a) It is not clear which prototype should be elicited in a context using frequency variations alone (without stories). Clearly, deviating from previous versions of representativeness, one would have to allow for conjunctive prototypes to explain the results for the Linda schools. Yet even conceding this, this would imply the prediction of a selection of conjunctions in *all* conditions where this conjunction is the most frequent of the four cells (cf. averaging heuristic); but this was not found in our results. In any case, in the present context, representativeness is underdetermined, at best, to explain the data. (b) Previously, evidence had been provided for differential sample-size effects not explicable by similarity measures used by Kahneman et al. (von Sydow, 2011; cf. von Sydow & Fiedler, 2012). (c) The standard model of representativeness does not naturally account for IFs, apart from CFs (cf. Bar-Hillel & Neter, 1993). (d) BL formulates a rational model of logical inclusion fallacies, whereas representativeness interprets CFs to be an irrational by-product of a useful heuristic (Kahneman & Tversky, 1996).

Despite such substantial differences, it remains open whether BL should be interpreted as providing a more precise, rational reformulation of a one application of representativeness or, perhaps more plausibly, as an alternative to both extensional probability and representativeness; but this hinges on the interpretation of "representativeness" that has been justly criticized as too vaguely defined (Gigerenzer, 1996; Gigerenzer, 2000; Wolford, 1991; Nilsson et al., 2008; cf. Kahneman & Frederick, 2005).

Overall, the corroborated system of logical inclusion fallacies based on simple contingency tables alone has not been predicted by any other main account of the CF.

Qualitative Theories

Although the experiments conducted did not focus on investigating the proposed extensional cues (for overviews, see Kahneman & Frederick, 2002, 2005; Barbey & Sloman, 2007), they support that even using several extensional cues simultaneously does not necessarily trigger extensional reasoning. In contrast, in the present tasks we used clear formulations of hypotheses, within-subject comparisons, no distractor hypotheses, clear nested sets, contingency tables with clear labels, and natural frequency information on all four cells of a contingency table; nevertheless, people appeared to remain in a non-extensional mode, since the goal of the task suggested judging the overall adequacy of alternative connectives. Thus extensional cues, even if used together, can be *in*sufficient to elicit extensional thought, even if enhancing such thought in other contexts.

These extensional cues nonetheless often facilitate extensional thought, particularly where an extensional solution is the only correct one. CFs may be based on either logically misunderstood formulations (e.g., Sides et al., 2002; Hilton, 1995; Messers & Griggs, 1993; Mellers et al., 2001; Hertwig et al., 2008; cf. von Sydow, 2014, 2015; but cf. Tentori, Bonini, & Osherson, 2004; Tentori & Crupi, 2012), a lack of frequency representation (e.g., Fiedler, 1988; Gigerenzer, 1991, 1996, 2000; Reeves & Lockhart, 1993; cf. Kleiter, 1994; Frederick & Kahneman, 2002, 2005; Zhu & Gigerenzer, 2006), between-subject designs (Tversky & Kahneman, 1983; Gigerenzer, 1996; Kahneman & Tversky, 1996; Mellers et al., 2001; Lagnado & Shanks, 2002), or unclear set inclusions (e.g., Sloman & Over, 2003; Sloman et al., 2003; Neace et al., 2008).

The present account does not deny that these context situations can matter (cf. our briefly mentioned Experiment 3), but the results suggest that the goal and interpretation of a probability judgment is not only determined by the presence of many discussed ‘extensional cues’.

Open Questions

First, despite the novel finding of systematic pattern-based IFs based on frequency information, it remains open whether a pattern-based approach can be generalized to traditional tasks about *individual* persons, such as Linda. Even though BL may use subjectively sampled (imagined) frequencies as input, there is only first evidence corroborating this application (von Sydow, 2013).

Second, apart from the tested model for *dyadic* logical relationships between *dichotomous* events, the role of alternative representations of ordinary-language connectives remains to be explored (Hertwig et. al., 2008). Von Sydow (2014) provides a BL account for conjunctions referring to patterns of marginal probabilities rather than one based on logical interaction patterns. Moreover, "and" often refers to the sum instead of the logical conjunction (von Sydow, 2015; Hertwig et al., 2008). Furthermore, polytomous category representations may play a role (von Sydow, 2015). This seems so even for the Linda example (as it involves, for instance, job categories, such as bank tellers, which are potentially polytomously represented).[8] The scope of the present paper, however, involves hypotheses formulations and 2×2 contingency matrices restricted to standard dichotomous, dyadic connectives. Nonetheless, variants of BL require elaboration in the future.

Third, our results show that discussed alternative theories of the CF cannot be universal accounts of CFs; yet I favour a polycausal theory of CFs (cf. Fiedler & von Sydow, 2015). Such a perspective requires research on boundary conditions of the different classes of CFs.

---

[8] Polytomous representations may play a role for all sorts of interpretations of ordinary language 'AND,' whether it is a dyadic, a marginal or a sum interpretation (cf. von Sydow, 2014, 2015).

Fourth, it is also an open question whether a simpler heuristics may implement pattern probabilities on an algorithmic level. BL formalises something like a similarity function between ideal logical patterns (or categories) and actual belief-patterns. A simpler ‘pattern heuristic’ might actually yield quite similar predictions. I am grateful to a reviewer’s suggestion of the following heuristic: Replace every frequency in the 2×2 data table that is greater than $N/4$ by a 1, and every frequency less than $N/4$ by a 0; then interpret the result as truth table (logical explanatory patterns supported by the data). In Experiment 1 this heuristic does actually lead to the same predictions as (deterministic) BL. It can likewise predict the first selections of Experiment 2, even though the second selections would require further specification. Because of its simplicity it seems to be the preferable algorithmic level account. However, it needs to be mentioned that this *specific* pattern heuristic does not seem to explain several other results obtained elsewhere: (a) it would predict an AND pattern for a [5 4 4 4] data pattern, where BL correctly predicted a tautology (von Sydow & Fiedler, 2012); (b) for a data pattern such as [14 27 19 0], the heuristics would predict a [0 1 1 0] hypothesis, but BL in line with the results predicts a [1 1 1 0] hypothesis (von Sydow, 2014, Panel E); and (c) this heuristic does not predict sample-size effects (von Sydow, 2011, Experiments 1a, 1b). Thus this specific pattern heuristic, albeit quite successful here, does not account for all findings. Nonetheless, in the future the basic intuition of a pattern account might be captured by a similar smart-and-simple heuristic that mimics pattern probabilities.

Finally, since extensional answers have been dominant in some contexts (Beach & Peterson, 1966; Pfeifer, 2013; Costello & Mathison, 2014), the different meanings of ‘probability’ should be investigated more directly, for instance by contrasting the goals of assessing (1) the probabilities of entities falling into a specific subset and (2) the adequacy of underlying rules.

Rational Inclusion 'Fallacies'?

Most researchers in this debate have regarded deviations from the inclusion rule (particularly of the conjunction rule) as a particularly startling and dramatic fallacy (Tversky & Kahneman, 1983; Kahneman & Frederick, 2002, 2005; Sides et al., 2002; Tentori et al., 2004; Mellers et al., 2001). In contrast, the pattern approach advocated here predicts a system of inclusion 'fallacies' as broadly rational if one considers the goal of determining which of several alternative noisy-logical connectives has most probably created a given data pattern. Although BL is inspired by Gigerenzer's (1996) critique of a context-blind application of (extensional) probability theory, it is also coherent with Kahneman and Tversky's (1996) seemingly opposed objections to 'normative agnosticism.' Going beyond a *direct* application of extensional probability theory (Kahneman & Tversky, 1996; Kahneman & Frederick, 2002; Kahneman & Frederick, 2005), BL is formulated as a rational proposal based on Kolmogorov's probability axioms, to attain a particular goal (Anderson, 1990). Whether BL is the only possible formalisation of pattern probabilities remains open. But note that BL, as mentioned in Table 1, does not abandon the axiom of additivity; it is only claimed that this axiom *need not* be applied at the same level as it is by extensional probability judgements. BL uses additivity on the level of the (second-order) probabilities of alternative noisy-logical rules, not on the level of extensions. If one gave up additivity altogether, a betting strategy could actually be implemented that would lead to inevitable gains or losses (Dutch Book argument, de Finetti, 1974; Gilio & Over, 2012, p. 124). However, it is assumed that with betting, adjudicators would evaluate the correctness of judgments in line with pattern probabilities; therefore no Dutch Book problem would arise.

As long as tasks request only probability for a dyadic logical relationship given some data, $P(A \text{ o } B|D)$, it is not specified whether one should use dyadic pattern probabilities—$P_H(A \text{ o } B|D)$—or standard extensional probabilities—$P_E(A \text{ o } B|D)$. Therefore participants commit no fallacy if they use pattern probabilities. In this case, the only fallacy involved would be that of the experimenter, adhering to a narrow norm regardless of the context. If the goal of the task concerns extensions, extensional probability should apply; but if it concerns the overall characterization of a logical interaction pattern, BL should apply. Although modifications might be needed to account for the intricacies of human probability judgments, it is hoped that BL constitutes another step in the intellectual co-evolution of developing systematic norms and investigating deviations from such norms (cf. Chater & Oaksford, 2000).

Conclusion

The present results support the idea of a system of pattern-based CFs, and their formalization by BL (von Sydow, 2011). Nonetheless, further tests are required covering other contexts (von Sydow & Fiedler, 2012; von Sydow, 2013, 2014, 2015). Moreover, although logical CFs are often assumed to form a monolithic class, it was suggested that there are presumably several causes of CFs (and, generally, of IFs). The current results merely provide substantial evidence for the claim that there is a consistent system of frequency-based IFs linked to logical patterns as yet not explicable by other theories. In the future, contrasts between pattern probabilities and extensional probabilities, universal and single-event predications, as well as between rational models and simpler heuristics need investigation in greater detail. Even if a simpler heuristic may account for results in the future, the present results strongly suggest this would require a kind of ‘pattern heuristic’.

List of References

Adams, E. W. (1986). On the logic of high probability. *Journal of Philosophical Logic, 15*, 255–279. DOI:10.1007/BF00248572

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

Bar-Hillel, M., & Neter, E. (1993). How alike is versus how likely is it: A disjunction fallacy in probability judgments. *Journal of Personality and Social Psychology, 65*, 1119–1131. DOI:10.1037/0022-3514.65.6.1119

Barbey, A. K., & Sloman, St. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences, 30*, 241–297. DOI:10.1017/S0140525X07001653

Beach, C. R., & Peterson, C. R. (1966). Subjective probabilities for unions of events. *Psychonomic Science, 5*, 307–308. DOI:10.3758/BF03328412

Betsch, T., & Fiedler, K. (1999). Understanding conjunction effects: The role of implicit mental models. *European Journal of Social Psychology, 29*, 75–93. DOI:10.1002/(SICI)1099-0992(199902)29:1<75::AID-EJSP916>3.0.CO;2-F

Biela, A. (1986). Psychological patterns in predicting disjunction and conjunction of clinical symptoms. *Acta Psychologica*, *61*, 183–195. DOI:10.1016/0001-6918(86)90080-6

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, *33*, 261–304. DOI:10.1177/0049124104268644

Busemeyer, J. R., Pothos, E., & Franco, R., Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment 'errors'. *Psychological Review, 118(2)*, 193–218. DOI:10.1037/a0022542

Carlson, B. W., & Yates, J. F. (1989). Disjunction errors in qualitative likelihood judgments. *Organizational Behavior and Human Decision Processes, 44,* 368–379. DOI:10.1016/0749-5978(89)90014-9

Chater, N., & Oaksford, M. (2000). The rational analysis of mind and behavior. *Synthese, 122*, 93–131. DOI:10.1023/A:1005272027245

Chater, N., & Oaksford, M. (Eds.). (2008). *The probabilistic mind: Prospects for Bayesian cognitive science*. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199216093.001.0001

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1–73. DOI:10.1016/0010-0277(95)00664-8

Costello, F. J. (2005). *A unified account of conjunction and disjunction fallacies in people's judgments of likelihood*. In B. Bara, L. Barsalou, & M. Bucciarelli (Eds.), Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society (pp. 494–499). Mahwah, NJ: Erlbaum.

Costello, F. J. (2009). Fallacies in probability judgments for conjunctions and disjunctions of everyday events. *Journal of Behavioral Decision Making, 22(3)*, 235–251. DOI:10.1002/bdm.623

Costello, F. J. & Mathison, T. (2014). *On fallacies and normative reasoning: When people's judgments follow probability theory*. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), Proceedings of the Thirty-Sixth Annual conference of the Cognitive Science Society (pp. 361–366). Austin, TX: Cognitive Science Society.

Crisp, A. K., & Feeney, A. (2009). Causal conjunction fallacies: The roles of causal strength and mental resources. *The Quarterly Journal of Experimental Psychology, 62(12)*, 2320–2337. DOI:10.1080/17470210902783638

Crupi, V., Fitelson, B., & Tentori, K. (2008). Probability, confirmation, and the conjunction fallacy. *Thinking & Reasoning, 14*, 182–199. DOI:10.1080/13546780701643406

de Finetti, B. (1974). Theory of probability (Vols. 1–2). (A. Machi & A. Smith, Trans.). Chichester, Wiley. (Original work published 1970)

De Neys, W. (2006). Automatic–heuristic and executive–analytic processing during reasoning: Chronometric and dual-task considerations. *Quarterly Journal of Experimental Psychology, 59*, 1070–1100. DOI:10.1080/02724980543000123

Doya, K., Ishii, S., Pouget, A., & Rao, R. (Eds.) (2007). Bayesian brain – Probabilistic approaches to neural coding. Cambridge, Mass.: MIT Press.

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences*, *7*, 454–459. DOI:10.1016/j.tics.2003.08.012

Evans, J. St. B. T., & Over, D. E. (2004). If. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780198525134.001.0001

Fantino, E., Kulik, J., Stolarz-Fantino, St., & Wright, W. (1997). The conjunction fallacy: A test of averaging hypotheses. *Psychonomic Bulletin & Review*, *4*, 96–101. DOI:10.3758/BF03210779

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research, 50*, 123–129. DOI:10.1007/BF00309212

Fiedler, K., & Juslin, P. (2006). Information sampling and adaptive cognition. Cambridge: Cambridge University Press.

Fiedler, K., & von Sydow, M. (2015). *Heuristics and biases: Beyond Tversky & Kahneman's (1974) judgment under uncertainty*. In M. W. Eysenck & D. Groome (Eds.), Cognitive psychology: Revisiting the classical studies (pp. 146−161). Los Angeles, London: Sage.

Fisk, J. E. (1996). The conjunction effect: Fallacy or Bayesian inference? *Organizational Behavior and Human Decision Processes, 67*, 76–90. DOI:10.1006/obhd.1996.0066

Fisk, J. E., & Slattery, R. (2005) Reasoning about conjunctive probabilistic concepts in childhood. *Canadian Journal of Experimental Psychology, 59*, 168–178. DOI:10.1037/h0087472

Fitelson, B. (2006). Inductive logic. In J. Pfeifer and S. Sarkar (Eds.), *Philosophy of Science:* An Encyclopedia (pp. 384–394). London: Routledge Press.

Foley, R. (2009). Beliefs, degrees of belief, and the Lockean thesis. In F. Huber, C. Schmidt-Petri (Eds.), Degrees of Belief (Synthese Library: Vol. 342), Heidelberg: Springer.

Gavanski, I., & Roskos-Ewoldsen, D. R. (1991). Representativeness and conjoint probability. *Journal of Personality and Social Psychology, 61*, 181–194. DOI:10.1037/0022-3514.61.2.181

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond 'heuristics and biases'. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology, 2* (pp. 83–115). Chichester: Wiley. DOI:10.1080/14792779143000033

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review, 103,* 592–596. DOI:10.1037/0033-295X.103.3.592

Gigerenzer, G. (1998). *Ecological intelligence: An adaptation for frequencies*. In D. D. Cummins & C. Allen (Eds.), The evolution of mind. New York: Oxford University Press.

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

Gigerenzer, G., Todd, P., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gilio, A., & Over. D. (2012). The psychology of inferring conditionals from disjunctions: A probabilistic study. *Journal of Mathematical Psychology*, *56*, 118–131. DOI:10.1016/j.jmp.2012.02.006

Gilovich, Th., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgement.* Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511808098

Girotto, V., & Gonzales, M. (2001). Solving probabilistic and statistical problems: A matter of information structure and question form. *Cognition, 78*, 247–276. DOI:10.1016/S0010-0277(00)00133-5

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive Science, 32*, 108–154. DOI:10.1080/03640210701802071

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51,* 334–384. DOI:10.1016/j.cogpsych.2005.05.004

Hahn, U., & Oaksford, M. (2007).The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review, 114*, 704-732. DOI:10.1037/0033-295X.114.3.704

Hájek, A. (2001). *Probability, logic, and probability logic*. In L. Goble (Ed.), The Blackwell Companion to Logic (pp. 362-384). Oxford: Blackwell.

Hartmann, S., & Meijs, W. (2012). Walter the banker – The conjunction fallacy reconsidered. *Synthese, 184*, 73–87. doi: 10.1007/s11229-009-9694-6. DOI:10.1007/s11229-009-9694-6

Hertwig, R., Benz, B., & Krauss, B. S. (2008). The conjunction fallacy and the many meanings of and. *Cognition, 108*, 740-753. DOI:10.1016/j.cognition.2008.06.008

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. DOI:10.1080/135467898394102

Hertwig, R., & Gigerenzer, G. (1999). The 'conjunction fallacy' revisited. How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12*, 275–305. DOI:10.1002/(SICI)1099-0771(199912)12:4<275::AID-BDM323>3.0.CO;2-M

Hertwig, R., & Volz, K. G. (2013). Abnormality, rationality, and sanity. *Trends in Cognitive Sciences, 17*, 547-549. DOI:10.1016/j.tics.2013.08.011

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin, 118,* 248–271. DOI:10.1037/0033-2909.118.2.248

Hintikka, J. (2004). A fallacious fallacy? *Synthese, 140*, 25–35. DOI:10.1023/B:SYNT.0000029938.17953.10

Jarvstad, A., & Hahn, U. (2011). Source reliability and the conjunction fallacy. *Cognitive Science 35*, 682–711. DOI:10.1111/j.1551-6709.2011.01170.x

Jenny, M. A., Rieskamp, J., & Nilsson, H. (2014). Inferring conjunctive probabilities from noisy samples: Evidence for the configural weighted average model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 203–217. DOI:10.1037/a0034261

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). A theory of meaning, pragmatics, and inference. *Psychological Review, 109*, 646–678. DOI:10.1037/0033-295X.109.4.646

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. DOI:10.1037/0033-295X.106.1.62

Kahneman, D., & Frederick, S. (2002). *Representativeness revised: Attribute substitution in intutitive judgment*. In Th. Gilovich, D. Griffin, & D. Kahneman (Eds.), Heurstics and biases: The psychology of intuitive judgement (pp. 49-81). Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511808098.004

Kahneman, D., & Frederick, S. (2005). *A model of heuristic judgment*. In K. J. Holyoak & R. G. Morris (Eds.), The Cambridge handbook of thinking and reasoning (pp. 267–293). Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*, 582–591. DOI:10.1037/0033-295X.103.3.582

Kao, S.-F., & Wassermann, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1363–1386. DOI:10.1037/0278-7393.19.6.1363

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review, 116(1)*, 20–58. DOI:10.1037/a0014282

Kern-Isberner, G. (2001). *Conditionals in nonmonotonic reasoning and belief revision*. Berlin, Heidelberg: Springer. DOI:10.1007/3-540-44600-1

Kleiter, G. (1994). *Natural sampling: Rationality without base rates.* In G.H. Fischer & D. Laming (Eds.), Contributions to mathematical psychology, psychometrics, and methodology (pp. 375–388). New York: Springer-Verlag. DOI:10.1007/978-1-4612-4308-3_27

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior, 36*, 210–226. DOI:10.3758/LB.36.3.210

Lagnado, D. A., & Shanks, D. R. (2002). Probability judgment in hierarchical learning: A conflict between predictiveness and coherence. *Cognition*, *93*, 81–112. DOI:10.1016/S0010-0277(01)00168-8

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology, 55,* 232–257. DOI:10.1016/j.cogpsych.2006.09.006

Maher, P. (2004). Bayesianism and irrelevant conjunction. *Philosophy of Science, 71*, 515–520. DOI:10.1086/423750

Mellers, B. A., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science, 12*, 269–275. DOI:10.1111/1467-9280.00350

Messer, W. S., & Griggs, R. A. (1993). Another look at Linda. *Bulletin of the Psychonomic Society, 31*, 193–196. DOI:10.3758/BF03337322

Navarro, D., Dry, M. J., & Lee, M. (2012). Sampling assumptions in inductive generalization. *Cognitive Science, 36(2)*, 187–223. DOI:10.1111/j.1551-6709.2011.01212.x

Neace, W. P., Michaud, St., Bolling, L., Deer, K., & Zecevic, L. (2008). Frequency formats, probability formats, or problem structure? A test of the nested-sets hypothesis in an extensional reasoning task. *Judgment and Decision Making, 3*, 140–152.

Nelson, J. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review, 112*, 979–999. DOI:10.1037/0033-295X.112.4.979

Nickerson, R. S. (2004). *Cognition and chance: The psychology of probabilistic reasoning*. Mahwah, NJ: Erlbaum.

Nilsson, H. (2008). Exploring the conjunction fallacy within a category learning framework. *Journal of Behavoral Decision Making, 21*, 471–490. DOI:10.1002/bdm.615

Nilsson, H., Juslin, P., & Olsson, H. (2008). Exemplars in the mist: The cognitive substrate of the representativeness heuristic. *Scandinavian Journal of Psychology, 49*, 201–212. DOI:10.1111/j.1467-9450.2008.00646.x

Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal influence. *Psychological Review, 111*, 455–485. DOI:10.1037/0033-295X.111.2.455

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608–631. DOI:10.1037/0033-295X.101.4.608

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780198524496.001.0001

Oberauer, K., Weidenfeld, A., & Fischer, K. (2007). What makes us believe a conditional? The roles of covariation and causality. *Thinking & Reasoning, 13*, 340–369. DOI:10.1080/13546780601035794

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, 11, 988–1010. DOI:10.3758/BF03196730

Over, D. E., & Evans, J. St. B. T. (2003). The probability of conditionals: The psychological evidence. *Mind & Language*, *18*, 340–358. DOI:10.1111/1468-0017.00231

Over, D. E., Hadjichristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, St. A. (2007). The probability of causal conditionals. *Cognitive Psychology, 54*, 62–97. DOI:10.1016/j.cogpsych.2006.05.002

Pfeifer, N. (2013). The new psychology of reasoning: A mental probability logical perspective. *Thinking & Reasoning, 19*(3-4), 329–345. DOI:10.1080/13546783.2013.838189

Reeves, T., & Lockhart, R. S. (1993). Distributional versus singular approaches to probability and errors in probabilistic reasoning. *Journal of Experimental Psychology: General*, 122, 207–226. DOI:10.1037/0096-3445.122.2.207

Schupbach, J. N. (2009). Is the conjunction fallacy tied to probabilistic confirmation? *Synthese, 84*, 13–27.

Schurz, G. (2001). Normische Gesetzeshypothesen und die wissenschaftsphilosophische Bedeutung des nichtmonotonen Schliessens. *Journal for General Philosophy of Science, 32*, 65–107. DOI:10.1023/A:1011209401546

Schurz, G. (2005). Non-monotonic reasoning from an evolutionary viewpoint: ontic, logical and cognitive foundations. *Synthese, 146*, 37–51. DOI:10.1007/s11229-005-9067-8

Sides, A., Osherson, D., Bonini, N., & Viale, R. (2002). On the reality of the conjunction fallacy. *Memory and Cognition, 30*, 191–198. DOI:10.3758/BF03195280

Sloman, St. A., & Over, D. (2003). *Probability judgement from the inside and out. In D. Over (Ed.), Evolution and the psychology of thinking: The debate* (pp. 145–169). Hove, New York: Psychology Press.

Sloman, St. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions. *Organizational Behavior and Human Processes, 91*, 296–309. DOI:10.1016/S0749-5978(03)00021-9

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*, 672–695. DOI:10.1037/0022-3514.94.4.672

Teigen, K. H. (1994). *Variants of Subjective Probabilities: Concepts, Norms, and Biases* (pp. 211-238). In: Wright, G. & Ayton, P. (Eds.). Subjective Probability. Chichester: John Wiley & Sons.

Tenenbaum, J., & Griffith, T. (2001a). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, *24*(4), 629–640. DOI:10.1017/S0140525X01000061

Tenenbaum, J., & Griffith, T. (2001b). *The rational basis of representativeness*. In J. Moore, & K. Stenning (Eds.), Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society (pp. 1036-1041). Austin, TX: Cognitive Science Society.

Tentori, K., Bonini, N., & Osherson, D. (2004). The conjunction fallacy: A misunderstanding about conjunction? *Cognitive Science, 28,* 467–477.DOI:10.1207/s15516709cog2803_8

Tentori, K. & Crupi, V. (2009). How the conjunction fallacy is tied to probabilistic confirmation: Some remarks on Schupbach (2009). *Synthese, 184*, 3–12. DOI:10.1007/s11229-009-9701-y

Tentori, K. & Crupi, V. (2012). On the conjunction fallacy and the meaning of and, yet again: A reply to Hertwig, Benz, and Krauss (2008). *Cognition, 122*, 123–134. DOI:10.1016/j.cognition.2011.09.002

Tentori, K., Crupi, V., & Russo, S. (2013). On the determinants of the conjunction fallacy: Probability versus inductive confirmation. *Journal of Experimental Psychology: General, 142*(1), 235–255. doi:10.1037/a0028770 DOI:10.1037/a0028770

Thüring, M. & Jungermann, H. (1990). The conjunction fallacy: Causality vs. event probability. *Journal of Behavioral Decision Making, 3,* 61–74. DOI:10.1002/bdm.3960030106

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131. DOI:10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315. DOI:10.1037/0033-295X.90.4.293

von Sydow, M. (2006). *Towards a Flexible Bayesian and Deontic Logic of Testing Descriptive and Prescriptive Rules: Explaining Content Effects in the Wason Selection Task* (Doctoral dissertation, University of Göttingen, Göttingen, Germany). Retrieved from ediss.uni-goettingen.de/handle/11858/00-1735-0000-0006-AC29-9.

von Sydow, M. (2007). *The Bayesian logic of the conjunction fallacy: Effects of probabilities and frequencies in contingency tables*. In D. S. McNamara, & J. G. Trafton (Eds.), Proceedings of the Twenty-Nineth Annual Conference of the Cognitive Science Society (pp. 1611–1616). Austin, TX: Cognitive Science Society.

von Sydow, M. (2009). *On a general Bayesian pattern logic of frequency-based logical inclusion fallacies.* In N.A. Taatgen & H. van Rijn (Eds.), 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*, 119–139. DOI:10.1016/j.jmp.2010.12.001

von Sydow, M. (2013). *Logical patterns in individual and general predication.* In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), Proceedings of the Thirty-Fifth Annual Conference of the Cognitive Science Society (pp. 3693–3698). Austin, TX: Cognitive Science Society.

von Sydow, M. (2014). *Is there a monadic as well as a dyadic Bayesian logic? Two logics explaining conjunction 'fallacies'.* In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), Proceedings of the Thirty-Sixth Annual Conference of the Cognitive Science Society (pp. 1712–1717). Austin, TX: Cognitive Science Society.

von Sydow, Momme (2014b). Bayesian Mental Models of Conditionals. *Cognitive Processing, 15* (Special Issue: Proceedings of the 12th Biannual conference of the German cognitive science society), 148-151

von Sydow, M. (2015). *Pattern probabilities for non-dichotomous events: A new rational contribution to the conjunction fallacy debate*. In Noelle, R. Dale, A. Warlaumont, J. Yoshimi, T. Matlock, C. Jennings, & P. Maglio (Eds.), Proceedings of the Thirty-Seventh Annual Conference of the Cognitive Science Society (pp. 2511-2516). Austin, TX: Cognitive Science Society. ISBN: 978-0-9911967-2-2 von Sydow, M. & Fiedler, K. (2012). *Bayesian logic and trial-by-trial learning*. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), Proceedings of the Thirty-Fourth Annual Conference of the Cognitive Science Society (pp. 1090–1095). Austin, TX: Cognitive Science Society.

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. DOI:10.1016/j.cognition.2007.08.003

Wolford, G. (1991). The conjunction fallacy? A reply to Bar-Hillel. *Memory & Cognition, 19*, 415–417. DOI:10.3758/BF03197147

Author Note

Momme von Sydow, Department of Psychology, Ruprecht-Karls-Universität Heidelberg, Germany.

Portions of the work were presented at the CogSci (Amsterdam, 2009). Further developments were presented at subsequent CogSci conferences (Sapporo, 2012; Berlin, 2013; Quebec, 2014; Passadena, 2015).

I would like to thank David Over, Klaus Oberauer, Valerie Thompson, and anonymous reviewers for valuable comments on earlier versions of the manuscript. I thank Anna Gast, Christina Botros, Johanna Frisch and Urszula Mihułowicz for conducting the studies and Martha Cunningham, Larry Fiddick, Johanna Frisch, York Hagmayer, Björn Meder, Ralf Mayrhofer, and Jonathan Nelson for helpful comments in an early phase of the manuscript. Finally, I am grateful to the Deutsche Forschungsgemeinschaft for a grant that made this work possible (DFG, Sy 111/2-1; part of the priority program ‘New Models of Rationality’, SPP 1516).

Tables

Table 1: Two types of probabilities, contrasting standard extensional probabilities, P_E, and intensional (pattern) probabilities, P_H, concerned with alternative generative pattern hypotheses

	Extensional probabilities	Intensional Probabilities
Normative answer to a question such as:	"How likely is it that an individual from a population is in a subset *H*, given *D*?"	"How likely is it that there is an underlying (probabilistic rule) *H* operating, given *D*?"
Application of Kolmogorov axioms	Direct application	Indirect application on the level of alternative hypotheses (second-order probabilities of explanatory probabilistic patterns)
Inclusion rule	$P(\text{subset}) \leq P(\text{superset})$	More specific (pattern) hypotheses can become more probable if the *pattern* more probably explains the evidence, given the data

Table 2: Numerical example for explanatory probability tables (PTs) for five logical hypotheses, assuming *r* = .5.

	A ∧ *B*		*A*		*B*		*A* >< *B*		*A* ∨ *B*	
	B	Non-*B*	*B*	Non-*B*	*B*	Non-*B*	*B*	Non-*B*	*B*	Non-*B*
A	.625	.125	.375	.375	.375	.125	.125	.375	.292	.292
Non-*A*	.125	.125	.125	.125	.375	.125	.375	.125	.292	.125

Table 3. Serial order of the schools (with cell frequencies) presented in Experiment 1

	First	Second	Third	Fourth
Order W	Maria	Linda	Magda	Xenia
	9,9,1,1	16,1,1,2	7,6,6,1	1,9,9,1
Order X	Linda	Xenia	Maria	Magda
	16,1,1,2	1,9,9,1	9,9,1,1	7,6,6,1
Order Y	Xenia	Magda	Linda	Maria
	1,9,9,1	7,6,6,1	16,1,1,2	9,9,1,1
Order Z	Magda	Maria	Xenia	Linda
	7,6,6,1	9,9,1,1	1,9,9,1	16,1,1,2

Table 4.

Correlations between models and empirical results (Experiments 1a and 1b)

	Probabilistic model	Deterministic model
BL (without / with log-odds), global	.98 / .82	.96 / .96
BL (without / with log-odds), detailed	.95 / .80	.93 / .93
Extensional probability, global	.42	.14
Extensional probability, detailed	.39	.13
Averaging heuristic, global	.62	.60
Averaging heuristic, detailed	.60	.57
Inverse probability, detailed	.37	.22
Confirmation_D, detailed	.42	.55
Confirmation_R, detailed	.37	.18

Note: Confirmation and inverse probability make different predictions for different presentation positions but do not provide clear predictions for the first position. BL, averaging heuristic, and extensional probability make predictions for all four positions. In order to provide a comparable measure, the correlations were calculated on the detailed level of each presentation position (apart from the first one) for all models (cf. main text).

Table 5.

Correlations between models and empirical results (Experiment 2)

	Hypothesis 1, probabilistic model	Hypothesis 1, deterministic model	Hypothesis 2, deterministic model
BL (without /with log-odds), global	.95 / .83	.93	.89
BL (without /with log-odds), detailed	.88 / .77	.86	.74
Extensional probability, global	.54	.13	.50
Extensional probability, detailed	.49	.14	.47
Averaging heuristic, global	.55	.38	.31
Averaging heuristic, detailed	.53	.36	.45
Inverse probability, detailed	.58	.36	.26
Confirmation_D, detailed	.61	.61	.40
Confirmation_R, detailed	.56	.36	.31

Note: The ‘detailed’ models are needed to model the order-dependent predictions of confirmation and inverse probability (cf. Experiment 1, Table 3).

Figures

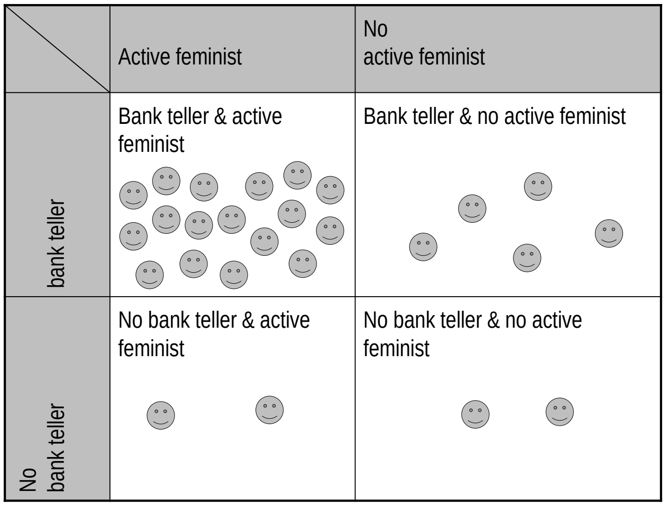


Figure 1. Contingency table with four cells (cell *a*: '*B* & *F*,' cell *b*: '*B* & non-*F*,' cell *c*: 'non-*B* & *F*,' cell *d*: 'non-*B* & non-*F*') and the labels used

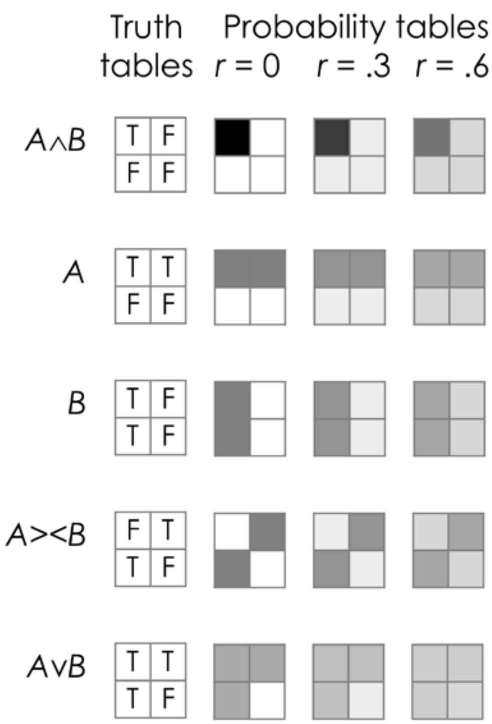
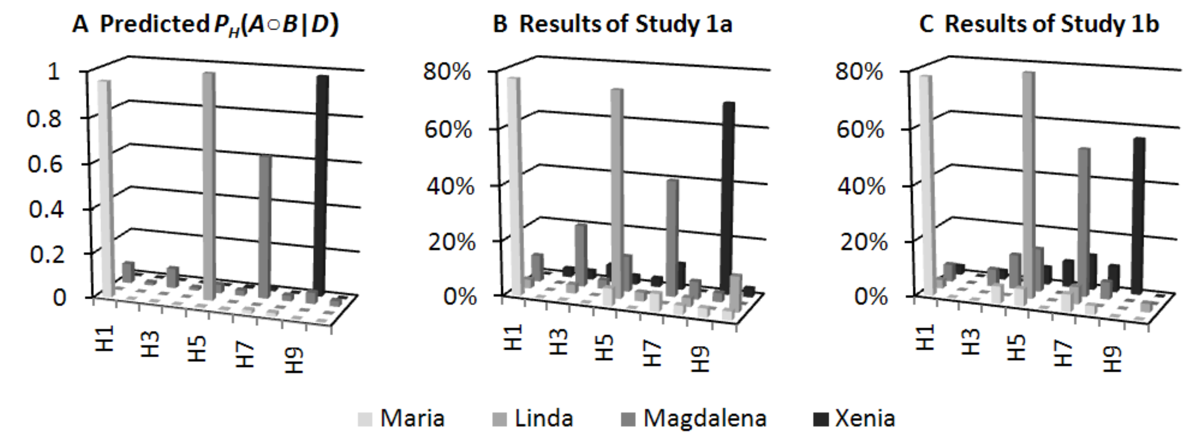


Figure 2. Illustrates the PTs of different connectives at different noise-levels *r*. The first column shows standard truth table definitions of the mentioned dyadic connectives (*T* = true; *F* = false), and the following three columns show PTs at different noise-levels (darker vs. brighter cells represent high vs. low probability).



D Selections predicted by extensional probability (EP) or Bayesian logic (BL)

Hypotheses "The graduates of this school generally become…"	Maria	Linda	Magda	Xenia
(H1) "…bank tellers, whether they are active feminists or not" (*A*)	BL, $CD_{2,3,4}$	CR_2		
(H2) "…no bank tellers, whether they are active feminists or not" ($\neg A$)			IN_4, CD_4, $CR_{3,4}$	IN_4, $CR_{3,4}$
(H3) "…active feminists, whether they are bank tellers or not" (*B*)		IN_2, CR_2	IN_3, CD_3	
(H4) "…no active feminists, whether they are active feminists or not" ($\neg B$)				
(H5) "…bank tellers AND at the same time active feminists" ($A \wedge B$)	IN_2	BL, IN, $CD_{3,4}$, CR	IN_3, CR_2	
(H6) "…bank tellers AND at the same time no active feminists" ($A \wedge \neg B$)	IN_4, IN_3, $CD_{3,4}$, CR		IN_2	IN_2, CR_2
(H7) "…bank tellers OR also active feminists or also both" ($A \vee B$)	EP		BL, EP	EP
(H8) "…bank tellers OR also no active feminists or both" ($A \vee \neg B$)	EP, $CD_{2,3}$			
(H9) "…EITHER bank tellers, who are no feminists, OR active feminists, who are no bank tellers" (*A* >-< *B*)			IN_2, $CD_{2,4}$	BL, $IN_{2,3}$, CD, CR_2
(H10) "IF the graduates of this school become bank tellers, THEN they become feminists" ($A \rightarrow B$)		EP, CD_2, IN_2	CD_3, IN_3	

Figure 3. Predictions and results of Experiment 1a (explicit concern with overall situation) and 1b (no explicit concern with overall situation): (A) Pattern probabilities (BL) for the observed samples in the Maria school <$n_a = 9$, $n_b = 9$, $n_c = 1$, $n_d = 1$> (first row), Linda school <16, 1, 1, 2>, Magdalena school <7, 6, 6, 1>, Xenia School <1, 9, 9, 1> (last row), assuming uniform priors for the ten hypotheses in question; (B) empirically found frequency of hypotheses selected 'most probable' in Experiments 1a and (C) 1b; and (D) hypothesis formulations used in Experiment 1a (cf. Experiment 1b) and predictions of BL, extensional probability (EP), inverse probability (IN), and confirmation (CD referring to difference-formulation, and CR to ratio-formulation). For all schools, the hypotheses with the highest values given a model (rounded) are indicated (if there are different predictions for different presentation positions, indices are used).

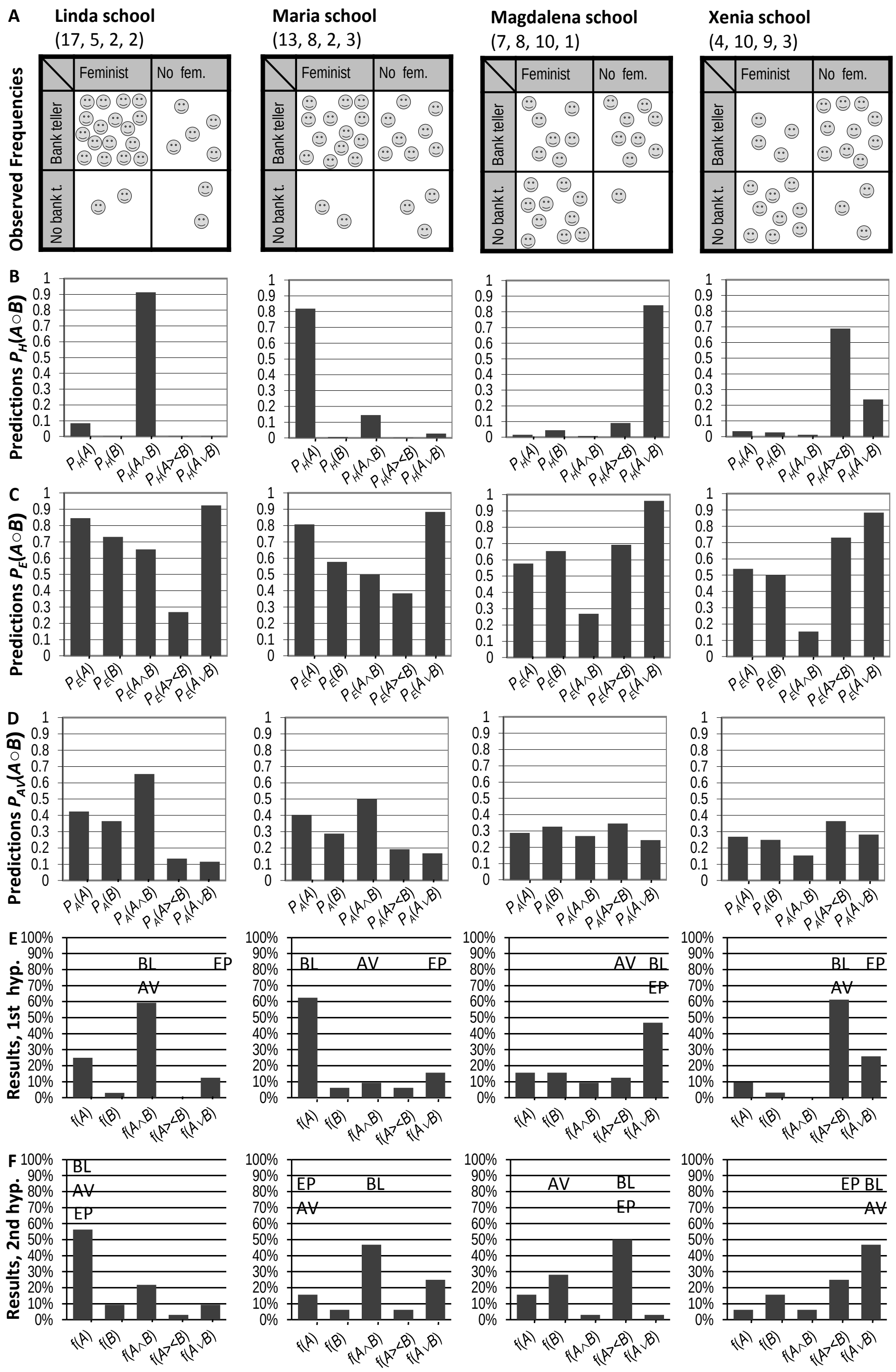
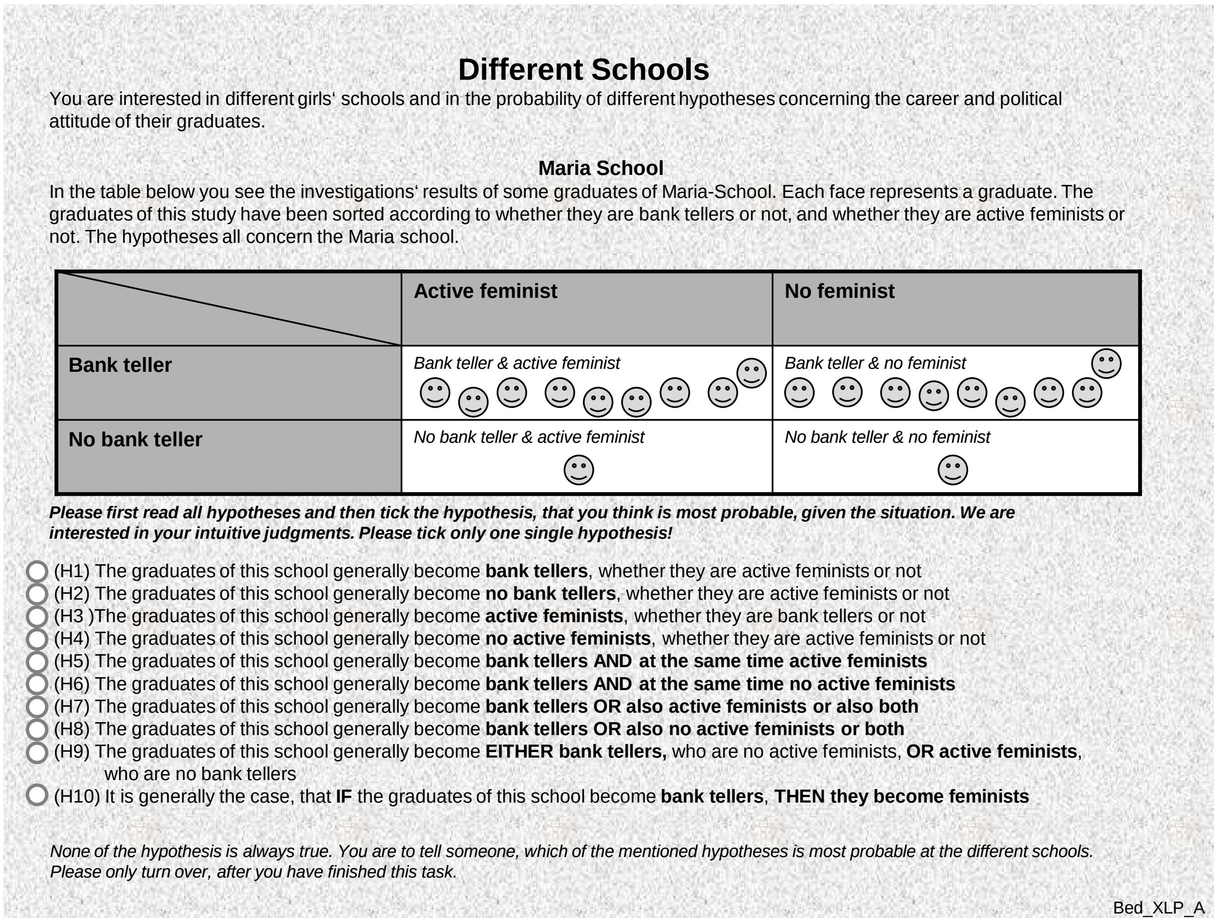


Figure 4. Predictions and the results of Experiment 2a, in which participants selected the two most probable hypotheses: (A) Frequencies presented; (B) predictions based on BL (without log-odds transformation); (C) extensional probabilities; (D) averaging heuristic; and (E, F) empirically found selections for the most and second-most probable hypotheses, with prediction of Bayes logic (BL), of extensional probability (EP), and with averaging heuristic (AV).

Appendix A



Different Schools

You are interested in different girls' schools and in the probability of different hypotheses concerning the career and political attitude of their graduates.

Maria School

In the table below you see the investigations' results of some graduates of Maria-School. Each face represents a graduate. The graduates of this study have been sorted according to whether they are bank tellers or not, and whether they are active feminists or not. The hypotheses all concern the Maria school.

	Active feminist	**No feminist**
Bank teller	*Bank teller & active feminist*	*Bank teller & no feminist*
No bank teller	*No bank teller & active feminist*	*No bank teller & no feminist*

Please first read all hypotheses and then tick the hypothesis, that you think is most probable, given the situation. We are interested in your intuitive judgments. Please tick only one single hypothesis!

- (H1) The graduates of this school generally become **bank tellers**, whether they are active feminists or not
- (H2) The graduates of this school generally become **no bank tellers**, whether they are active feminists or not
- (H3)The graduates of this school generally become **active feminists**, whether they are bank tellers or not
- (H4) The graduates of this school generally become **no active feminists**, whether they are active feminists or not
- (H5) The graduates of this school generally become **bank tellers AND at the same time active feminists**
- (H6) The graduates of this school generally become **bank tellers AND at the same time no active feminists**
- (H7) The graduates of this school generally become **bank tellers OR also active feminists or also both**
- (H8) The graduates of this school generally become **bank tellers OR also no active feminists or both**
- (H9) The graduates of this school generally become **EITHER bank tellers,** who are no active feminists, **OR active feminists**, who are no bank tellers
- (H10) It is generally the case, that **IF** the graduates of this school become **bank tellers**, **THEN they become feminists**

None of the hypothesis is always true. You are to tell someone, which of the mentioned hypotheses is most probable at the different schools. Please only turn over, after you have finished this task.

Bed_XLP_A

Figure A1. Translated task instruction used in Experiment1a (Maria school).