# SONDERFORSCHUNGSBEREICH 504

Rationalitätskonzepte,
Entscheidungsverhalten und
ökonomische Modellierung

No. 07-54

**Pseudocontingencies - A key paradigm for understanding adaptive cognition**

Klaus Fiedler*

July 2007

Thanks are due to Nick Chater, Peter Freytag, Ralph Hertwig, Florian Kutzner, Mike Oaksford, and Tobias Vogel for their thoughtful and constructive comments on an earlier draft of this chapter. Correspondence concerning should be addressed to Klaus Fiedler, Psychologisches Institut, Universität Heidelberg, Hauptstrasse 47-51, 69117 Heidelberg, Germany. Fax: + 49 / 6221 / 547745. Email: kf@psychologie.uni-heidelberg.de

*Sonderforschungsbereich 504/ Universität Heidelberg, email: Klaus_Fiedler@psi-sv2.psi.uni-heidelberg.de

Universität Mannheim
L 13,15
68131 Mannheim

Pseudocontingencies – A key paradigm for understanding adaptive cognition

Klaus Fiedler (University of Heidelberg)

To appear in:

N. Chater & M. Oaksford (Eds.)

*The probabilistic mind:*

*Prospects for rational models of cognition*

*Running Head*: Pseudocontingencies – A key paradigm

*Author Note*: Thanks are due to Nick Chater, Peter Freytag, Ralph Hertwig, Florian Kutzner, Mike Oaksford, and Tobias Vogel for their thoughtful and constructive comments on an earlier draft of this chapter. Correspondence concerning should be addressed to Klaus Fiedler, Psychologisches Institut, Universität Heidelberg, Hauptstrasse 47-51, 69117 Heidelberg, Germany. Fax: + 49 / 6221 / 547745. Email: kf@psychologie.uni-heidelberg.de

The probabilistic mind has to mirror the structure of the probabilistic world. Because the mind reflects the environment, the topic of the present volume should not be misunderstood as referring only to intrapsychic processes within individual organisms' brains or minds. Rather, the "probabilistic mind" refers to the adaptive match between cognitive functions and environmental tasks and affordances. Studying the probabilistic mind calls for a cognitive-ecological approach that relates mental functions to environmental structures, rather than a purely cognitive approach that relates individual mental functions to micro-level intrapsychic processes, such as neuronal processes. In this respect, the cognitive-ecological perspective that guides the present article, and most other articles in the present book, may be conceived as complementary to a neuro-scientific approach to the human mind. The cognitive illusion that is in the focus of the present chapter – called pseudocontingencies – highlights the need to study the top-down constraints imposed by the environment on cognitive behavior, which are quite distinct from the bottom-up constraints of internal neuronal processes.

*The Nature of the Probabilistic World*

Before I can explain and illustrate the key concept of pseudocontingencies, a moment of reflection is in order about the nature of the probabilistic world. What renders nature so uncertain and so difficult to handle? A most common answer suggests, like the title of this book, that difficulty arises because the world is probabilistic, rather than deterministic. Real correlations are hardly ever perfect. Although there is no question that imperfect, probabilistic correlations are more difficult to represent in memory than deterministic relations, I believe that this idea provides only an impoverished picture of the actual vicissitudes of the complex world. Imperfect, merely probabilistic relations between environmental variables need not in and of themselves be taxing and complicating. They can be quite plausible, natural, and they can create optimism. That the relationship between socio-economic status and income is less than perfect creates hope and chances in those belonging to the lower social class. Any optimism presupposes that the future world is not totally determined. Thus, probabilistic relations are desired properties in a world that should be predictable on one hand while leaving latitude for change and improvement on the other hand.

What is really bothersome and a permanent source of conflict and erroneous decisions, though, is the fact that the "true" relationship that actually holds between any two variables is often ambiguous, or indeterminate, because there is more than one correct or best solution. To illustrate this ultimate source of uncertainty, which pervades virtually all normative models of rational behavior, let us look for the common denominator underlying the following paradigms of psychological research:

(a) A *delay of gratification* task (Metcalf & Mischel, 1999) involves a forced choice between one option leading to a relative short-term advantage (e.g., shorter education → earlier job with a reasonable income) and another option leading to a relative long-term advantage (longer education → qualification for greater variety of jobs). Determining the "best" option involves a trade-off between short-term and long-term utilities, and a decision for the most appropriate time frame to assess the utility. There is no a priori principle saying that a long-term frame is more "real" or "more rational" than a short-term frame.

(b) A *dilemma* task is by definition a task that involves a conflict between two strategies, to defect or to cooperate. The pay-off of defecting is higher at the level of individual trials. However, averaging across many trials of a dilemma game, cooperation is the more successful strategy, because extended defection evokes negative payoffs or sanctions from the environment that override the seeming advantage. What strategy is optimal cannot be determined absolutely. It depends on the level of analysis, as evident in economists' differential treatment of single-shot games and games repeated over multiple trials.

(c) Many *optimizing* problems call for a choice between two options. An animal whose major adaptive task is to find and collect food may experience that the average amount of food is higher in location A than in location B. At the level of the individual animal, then, it is rational to move to A. However, when aggregating over many individuals of the same species, this implies that other animals will move to A as well, causing a hard and dangerous competition for resources and a virtual decrease in the actual amount of food provided well below the average amount that could be expected from the less prominent location B.

(d) Conversely, *Simpson's (1951) paradox* typically starts with an overall correlation showing that, say, more female than male applicants for graduate studies are rejected. However, as this aggregate relationship is broken down to a lower level of analysis, the apparent correlation turns out to be spurious. Within both of two graduate programs, females are more successful than males. The apparent disadvantage of females in the overall analysis merely reflects the fact that most females apply to the more difficult graduate program with a greatly enhanced rejection rate. In other words, when the impact of the unequal graduate programs is partialled out (i.e., when changing from an overall to a more specific level of analysis), the sign of the observed correlation is reversed. Again, there is no a priori basis for considering the partial correlation more correct than the zero-order correlation. To be sure, it is possible that a true female advantage is only visible when the impact of specific programs is controlled for. However, it is also possible that the higher rejection rate of the seemingly more difficult program merely reflects a higher rate of female applicants.

(e) When it comes to *correlations over time*, spectral analysis or time series analysis tells us that the correlation that holds between two variables over time depends on the frequency or periodic unit. Sometimes, correlations emerge strongly when considering time segments of seconds or milliseconds (e.g., EEG data) but disappear when aggregating measures over larger time units. Other trends of correlations can only be assessed at the level of long-term moving averages (e.g., the global warming effect and its correlates). In still other domains (e.g., the stock market), correlations (e.g., between share values and unemployment rates) may be positive in the short run and negative in the long run. In general, time-series analyses highlight the fact that different frequency filters render different phenomena visible.

(f) Last but not least, to add a prominent example from social psychology, the *hidden-profile* paradigm in group decision making (Mojzisch & Schulz-Hardt, 2006; Stasser & Titus, 1985) involves divergent decision preferences at the levels of the entire group and its individual members. One option excels in the information available to individual decision makers; applying an "optimal" majority or Condorcet rule (Hastie & Kameda, 2005) will lead the group to choose this very option. However, when all information distributed over all group members is shared by communication, another option may be superior. Although the common premise in this paradigm is that the group-level information is the validity criterion, there is no rational basis for this assumption. It is very possible that the quality of the totally shared information is worse than the individual-level information. In any case, correlations and preference structures can change and even reverse when information is aggregated over individuals or group members – a huge challenge for all democratic societies. Thomas Schelling’s (1978) book on micromotives and macrobehavior anticipated these intriguing insights three decades ago.

*Ecological Correlations Provide a Statistical Model*

All these paradigms share, as a common denominator, the disillusioning insight that globally correct solutions for these puzzles and pitfalls of the probabilistic world may not exist. What is correct, rational, or beneficial can only be determined locally, that is, conditional on pragmatic assumptions that specify a specific perspective, aggregation level, or units of analysis. Standard normative models, such as correlation statistics or Bayesian calculus, only afford a locally rational solution, once a specific perspective and level of analysis has been chosen. They offer no way of dealing with the trade-off between the solutions pertinent to different aggregation levels.

For a statistical model of the generic structure underlying these multi-level problems, let us refer back to the old notion of ecological correlations, which provides a starting point for our recent research on pseudocontingencies, the focus of the present chapter. As indicated by

Robinson (1950) and explained statistically by Hammond (1973), the correlation between race and illiteracy can be close to zero when computed across individuals. However, when computed at the level of districts or larger ecologies, the correlation between the average rate of Black people and the average illiteracy rate can rise to over +.90. In a similar vein, the correlation between price and quality can vary greatly when computed over either individual consumer products or markets or providers. Or, the relation between socio-economic status and academic performance can be quite different when considering either individual students, or entire classes, or school systems.

One must not discard these examples as simply reflecting reliability artefacts (i.e., the enhanced reliability of aggregate units of measurement). Indeed, it is easy to find correlations that are stronger at individual than at aggregate levels. The actual reason for divergent correlations is that different causal factors can be operating at different levels. Consider the following example, which provides a generative model to understand and simulate the degree of divergence between aggregation levels that is possible. Imagine there are 50 towns in a country, differing in the tourist ratio (relative to the total population of a town) and the average consumption rate (i.e., the amount of money spent by an average person on a day). Across all towns, the correlation between tourism and consumption is probably very high, because nice towns attract both tourists and rich people whereas nasty towns will remain for poor people and have few tourists. In contrast, assuming that all residents of the country have a clearly higher income than tourists from other countries (if the focal country is, say, Switzerland), the individual correlation between tourism and consumption (within towns) may be negative. That is, the higher the rate of tourists (with markedly lower income) in any town, the lower the consumption. Thus, a causal parameter of towns (i.e., attractiveness) can account for a high positive correlation, while a causal parameter for individuals (i.e., income) can account for the co-existing negative correlation. No artefact is involved. The two correlations are equally correct. They just reflect a genuine divergence between aggregation levels. Using this problem for recent simulations and decision experiments, it was easily possible to create co-existing correlations as positive as +.76 at town level and, yet, as negative as –.48 at individual level.[1]

Problems like these are neither artefactual nor far-fetched. In many real-world domains, they appear to be the rule rather than the exception. In psychological research, for instance – to put the finger on a nearby-ecology – researchers use to employ group data to make inferences about genuinely individual processes (such as memory or emotions). There is no guarantee, however, even in experimental research, that group aggregates reflect the same relationships that exist within individuals and that are often the focus of theoretical interest.[2] Researchers who commit the category mistake to base inferences about individual processes on group averages come very close to the pseudocontingency illusion to be introduced next.

*The Pseudocontingency Illusion – A Cognitive Analog of Ecological Bias*

For a more vivid illustration of this cognitive illusion, which can be understood as a cognitive analog of Robinson's (1950) ecological bias, consider a teacher who is confronted with the task of evaluating and grading the performance of boys and girls in a physics class (cf. Figure 1). Imagine a teacher who, at the beginning of a new school year, enters a class that has a high baserate of boys (75%) and a high baserate of high achievement (75% correct responses). In another class, then, the teacher encounters a low baserate of boys (25%) jointly with a low baserate of good achievement (25%). Empathizing with the teacher, we understand that at this point she will already assume a positive correlation between male gender and achievement in physics. This conviction will increase to certainty when there are two other classes, again one with high baserates and one with low baserates of both attributes. However,

[1] To simulate *n* individuals' consumption, one only has to use a (e.g., normally distributed) random variable of inter-individual consumption differences and add a salary parameter *s* for residents (rather than tourists) and to add an attractiveness parameter *a* for all people (residents as well as tourists) in attractive towns. Depending on the value of *a* and *s*, relative to the variance between individuals, the resulting correlations can differ markedly.
[2] Although experimental designs based on randomized groups attempt to eliminate the systematic variance between groups, the problem may still persist in more subtle ways, for instance, when experimental treatments (e.g., emotion treatments) applied to groups do not guarantee the same influence on every individual.

a glance at individual students' performance shows that within all four classes, the good-achievement rate is lower for boys than for girls (see Figure 1). Pooling across all four classes, the correlation turns out to be zero. A teacher who – like the empathic reader of this paragraph – believes to have experienced an advantage of boys although boys are in fact not superior, or even inferior, to girls, has fallen prey to the pseudocontingency (PC) illusion.[3]

*Definition of the PC illusion*. How can the PC illusion be defined and explicated more precisely? – To introduce the concept, consider the elementary case of a relation between two variables, X and Y, in a two-dimensional space. (A more general definition extends to an *n*-dimensional relation in *n*-dimensional space.) To keep within the preceding example, let X and Y be two dichotomous variables, student gender and achievement. The genuine contingency between these two variables is determined by the 2 x 2 joint frequencies of a contingency table (cf. Figure 2). Virtually all previous research assumes that the cognitive process of contingency assessment is a function of the four stimulus frequencies, *a, b, c, d*, in accordance with standard statistical correlation models (Allan, 1990; Alloy & Tabachnik, 1984; Fiedler, 2000a; McKenzie, 1994). Errors and biases in correlation assessment are attributed to unequal attention and differential weights given to these four cells, due to prior expectancies, the salience of variable levels, or the asymmetry of present versus absent information. In any case, it is presupposed that human (like animal) contingency assessment is based on a cognitive function that uses the joint frequencies, or cell entries, as its argument.

In contrast, a pseudocontingency (PC) is an inference rule that uses the marginals of the contingency table, rather than the cell entries (Fiedler & Freytag, 2004; Fiedler, Freytag & Unkelbach, 2007; Freytag, 2003). In other words, the PC algorithm (mis)takes two skewed baserate distributions for a contingency. When the marginals or baserate distributions are skewed in the same direction (i.e., mostly male students and mostly good achievement), the

[3] Note that the term "illusion" does not imply the violation of an incontestable norm of rationality. PC illusions can be functional or dysfunctional, depending on what level of aggregation is adequate, just as the functionality of other illusions, like overconfidence, depends on the learning environment (cf. Haselton & Funder, in press; Hoffrage, Hertwig & Gigerenzer, 2000).

inferred contingency between male gender and achievement is positive. When the marginals or baserate distributions are skewed in opposite directions (i.e., mostly male students but rarely good achievement), the inferred contingency is negative. As vividly shown in Figure 2, inferring a contingency from the alignment of two baserate distributions is not justified, because the same baserates allow for positive, zero, and positive correlations. Confusing baserates with contingencies is like confusing two main effects (i.e., a row difference and a column difference) with an interaction (changing column differences as a function of rows). However, judges and decision makers – or more generally: organisms – commit this category mistake in many different task contexts, as evident from a good deal of empirical evidence reviewed in the next section.

In fact, the PC illusion is not as stupid as it may appear at first sight. Like Robinson's (1950) ecological bias and the other multi-level problems depicted at the outset, the PC illusion produces an error at one level but a sound inference at another, aggregate level. After all, at the level of classes, the rates of boys and of good achievement are jointly elevated, in comparison to some normative standard that usually holds for other classes. Indeed, by exposing the teacher to a contrast class with a low baserate of boys and a low baserate of good achievement (regardless of the within-class correlation across students), the teacher's PC illusion could be amplified. However, such an explicit ecological correlation between the proportions of boys and higher achievers across two or more classes or ecologies is not strictly necessary for the PC effect to occur. Even if there is but one class or ecology, the teacher can use her prior knowledge of normal classes to infer the covariation of baserates across ecologies, whether explicitly observed or implicitly memorized.

Thus, to complete the definition, PCs result when the correlation of category baserates is (mis)taken for inferring the correlation of individual measures. The term PC refers to illusions arising from this inference rule; it does not refer to the erronesous outcome of an illusory correlation inference, which can reflect many other processes (cf. Fiedler, 2000a). The PC illusion occurs under many conditions that render the efficient assessment and encoding of aggregate-level information (i.e., baserates) more likely than individuating information (i.e., joint frequencies). By analogy, a generalized definition of PCs in *n*-dimensional space says that inferences on complex contingencies involving *n* dimensions are often based on observations gathered in an aggregate space of lower dimensionality (resulting from aggregation over some dimensions). Thus, with reference to the above PC example, cognitive inferences about a three-dimensional data array, involving student performance x gender groups x students within gender groups, are based on a two-dimensional array involving aggregate scores for performance x gender groups.[4]

*Empirical Evidence*. A cursory review of empirical evidence for experimentally controlled PC effects will further help to illustrate the various manifestations of the illusion. Note that, psychologically, PCs suggest a tendency for higher-order, aggregate correlations to dominate and overshadow lower-order, individuating correlations. A recent series of experiments conducted within a simulated classroom paradigm (Fiedler et al., 2007) speaks to the very example that was used here to introduce the phenomenon, namely the correlation between student gender and achievement.

In this paradigm, participants are asked to take the role of a teacher who has to observe the performance of a class of 16 students, 8 boys and 8 girls, represented graphically on the computer screen. Each lesson is devoted to a particular subject matter, such as maths, physics, English or German. Over an extended period of time, the teacher can select a knowledge question from a pull-down menu of questions representing the subject matter. Once a question is announced, a subset of all students raises their hand, and the teacher selects one student who then provides either a correct or a wrong answer. Across many question-answer cycles of this kind, the teacher can assess the achievement of all 16 students in the class. As each

[4] More generally, PC-like inferences occur whenever a higher-dimensional problem design (e.g., a 4-dimensional design involving factors A x B x C x D) is "studied", either in people's mind or in science, through one or more sub-designs (e.g., design A x B; design C x D; design A x D etc.), which aggregate over the levels of the omitted factors.

student’s true ability parameter (i.e., his or her probability of providing a correct response) and motivation parameter (i.e., his or her probability of raising hand) are controlled by the computer program that drives the simulated classroom, both the accuracy and the potential biases in the teachers’ assessment can be studied systematically.

In one experiment, teachers were asked to test the hypothesis that boys are good in science (maths and physics) whereas girls are good in language, corresponding to common gender stereotypes. This led most participants to engage in positive testing (Klayman & Ha, 1987; Oaksford & Chater, 1994), that is, to ask more questions to boys in science and to girls in language lessons. Consequently, the gender baserate distributions were skewed in opposite directions for science and language lessons; there were clearly more answers from boys in science but clearly more answers from girls in language. Distinct PC effects were induced when these skewed gender baserates were aligned with the skewed correctness baserates of smart students (with a correctness rate of 80%). For smart students in science, the coincidence of mostly male responses and mostly correct responses led teachers to judge the ability of smart boys higher than the ability of smart girls (with the same objective ability parameter). For language lessons in contrast, mostly female responses and mostly correct responses led teachers to judge smart girls higher than (objectively equivalent) smart boys. Closer analyses revealed that this finding was confined to those teachers who actually engaged in positive testing (i.e., who actually produced skewed gender distributions).

That the PC bias reflects the alignment of skewed baserates, rather than expectancies based on gender stereotypes, was demonstrated by the reverse task instruction, namely, to test the hypothesis that (in this particular class) girls tend to be good in science but boys tend to be good in language. Positive testing now led teachers to mainly focus on girls in science and on boys in language, thus producing an opposite skew in the gender baserates. As a consequence, mostly female and mostly correct responses led teachers to judge smart girls higher than smart boys in science. In language, in contrast, mostly male responses together with mostly correct responses led smart boys appear superior to smart girls. Again, the biases were confined to those teachers who actually engaged in positive testing, the precondition of skewed gender baserates.

In another experiment from the same series, a PC effect accounts for the impact of the class context on the evaluation of individual students' performance. In one class, the ability of all students was set to a constant correctness baserate of 70%. In another class, the correctness baserate was constantly low, 30%. Within both ecologies, the individual students' motivation parameters varied from 20% to 50% and 80%. Thus, the true correlation between students' motivation and their ability was by definition zero, because individual ability was invariant and the correctness of responses to specific questions depended on the computer's random generator, which is independent of whether a student had raised his or her hand or not.

Nevertheless, distinct PC effects reflected subjectively inferred correlations between motivation and ability. In a high-ability class environment, with a high correctness baserate, the motivation baserates for highly motivated students was skewed in the same direction, suggesting a positive PC, which led teachers to judge the ability of high-motivation students higher than low-motivation students. In contrast, in a low-ability environment, the low correctness baserates were skewed in a direction opposite to the high hand-raising baserates of highly motivated students. The resulting negative PC suggested a negative relation between motivation and ability, leading teachers to judge the ability of high-motivation students lower than the ability of low-motivation students (whose low motivation baserates were well aligned with the low correctness baserates).

In still other experiments, PC effects demonstrated the impact of group aggregates on judgments of individual students. The class was divided into two subgroups of eight students supposed to come from different former classes or teachers. In one subgroup, there were mostly high-ability students and high-motivation students, whereas the other subgroup consisted of mostly low-ability and low-motivation students. However, crucially, the correlation between ability and motivation at the level of students was zero, as the ratio of high to low ability students was the same among both high and low motivation students. Nevertheless, when teachers rated the individual students' ability and motivation at the end of the session, the resulting sets of 16 ratings were correlated, reflecting a typical PC effect. The coincidence of high baserates of both attributes in one subgroup and low baserates of both attributes in the other subgroup – that is, the existing correlation between ability and motivation baserates at the level of subgroups – misled teachers to infer a corresponding correlation at the level of individual students.

An analogous finding was obtained in still another experiment between individual students' positions on two political attitude topics, as uttered in a civics lesson. Although the correlation between the 16 students' pro and con positions on one attitude were completely uncorrelated with their pro and con stands on the other attitude, the teachers believed to have seen a correlation because one subgroup of students held mostly pro attitudes on both topics, whereas another subgroup held mostly con attitudes on both topics. The sign of the PC illusion was reversed, that is, teachers believed that pro positions on one attitude came along with con positions on the other attitude, when the baserates of pros and cons in the two subgroups were skewed in opposite directions.

Convergent evidence for PC illusions that reflect the same theoretical principle (alignment of skewed baserates) comes from a whole variety of task settings and content domains. Conceptual replications include PCs between individual scores on different personality tests, when respondents belong to different groups with different baserates of test scores (Fiedler & Freytag, 2003); PCs between dieting and symptoms of patients in two wards of a hospital (Fiedler & Freytag, 2004); PCs between a couple's responses to the items of a partner questionnaire when several subtests yield different baserates of yes and no responses (Freytag, Fiedler, Randoll & Vogel, 2007); between the occurrence of a virus and a disease in different geographical areas (Fiedler & Graf, 1990); or between the desirability of behavior and the belongingness to one of two social groups with different towns serving as ecologies (Meiser, 2006; Meiser & Hewstone, 2004).

In more recent studies, we were even able to demonstrate PC effects in sequential learning and speeded classification tasks such as evaluative priming with different baserates of positive and negative primes and targets (Fiedler, Blümke & Unkelbach, 2007), in the Implicit Association Test (IAT) with different baserates of target attributes and valence attributes (Blümke & Fiedler, 2007), and in Goodie and Fantino’s (1996) probability-learning paradigm (Kutzner, Freytag, Vogel & Fiedler, 2007).

Of particular interest is the analysis of the specific task conditions that give rise to PC illusions. An overview of the available evidence suggests, first of all, that the phenomenon generalizes over a variety of conditions. PC effects have been shown to result from the alignment of skewed baserate distributions in a single group, in two groups, or in four groups or categories. PCs occur whether the groups or ecologies can be assumed to reflect a common cause of the skewed baserates (i.e., preceding therapy in one group as a cause of skewed test baserates) or a common effect (i.e., therapy as a consequence of observed test values).

*Setting PCs apart from genuine contingencies*. Most importantly, the illusion generalizes over different presentation modes, called successive versus simultaneous. In the successive presentation mode, participants are first presented information about individuals’ high versus low values on one variable (e.g., test X) in one run, before they are later presented information about a second variable (test Y) in another run. In other words, they are not fed with genuine contingency information about the joint occurrence of X and Y in the same persons. Rather, they merely receive information about the uni-variate distribution of each variable within the group. It is this condition that clearly sets PCs apart from the usual contingency assessment paradigm, in which the stimuli are always bi-variate observations of both variables shown at the same time. Thus, in the successive mode, participants have no chance to solicit the contingency proper; the only remarkable finding is that participants readily infer subjective contingencies from two separate series of uni-variate observations.

In the simultaneous presentation mode, in contrast, joint observations for both variables (e.g., test X and Y) are presented simultaneously, linked to the same person, thus providing all information that is necessary to assess the genuine contingency. PCs are pitted against contingencies (cf. Figure 2a); that is, skewed marginal distributions suggest a PC opposite to the contingencies given by the cell entries of the contingency table. It is remarkable that even in this “home domain” of contingency assessment, PCs often override contingencies proper. In other words, even though the joint frequencies or cell entries are available, participants utilize the baserates or marginals for contingency inferences. This intrusion of PC illusions into the contingency domain suggests the challenging idea that many previous findings on illusory correlations may, to an unknown degree, reflect hidden PC effects.

*Learning Environments Fostering the Evolution of PC Illusions*

Why should evolution have allowed homo sapiens to develop such a serious category mistake, given the great adaptive value of accurate contingency assessment? Why should an organism exposed to the contingency in Figure 2a, which is negative ($r = -.20$), make predictions from individual X to Y scores as if the relation were positive, as suggested from the alignment of skewed distributions (mostly high values on both X and Y)?

Upon some reflection, there are indeed several good reasons for PCs. An analysis of the learning environments in which organisms typically have to assess contingencies shows that PC-based inferences are not at all stupid or irrational. First of all, it has to be kept in mind that PCs are not simply wrong or fully detached from reality; rather, they correctly reflect ecological correlations that hold at an aggregate level of groups or higher-order categories. The question then becomes why and under what conditions is homo sapiens inclined to assess ecological correlations at aggregate level rather than individuating correlations at more specific levels, even when a decision problem calls for individuating information?

*Nasty environment for contingency assessment*. A simple and striking answer to this crucial question can be found if one considers the structure of the probabilistic world surrounding the probabilistic mind. To illustrate this point, let us return to the teacher who is to learn correlates of student achievement. There are many potential correlates in the information environment: student motivation, personality of teacher, instruction style, socio-economic status, TV consumption, variation between subject matters, and so forth (cf. Figure 3). As the teacher gathers data about student achievement, she does not know which particular correlate will be the focus of a judgment problem at some future time. To be prepared for any problem (i.e., relating achievement to any of these correlates), the teacher would have to assess the full multivariate contingency table. Here, however, she encounters a number of insurmountable problems. First, the environment rarely provides us with complete multivariate data points. At the very time when a student's achievement is observed, the corresponding data for many other variables (SES, former teacher's method, parents' style) may not be available. Second, even if it were available, the teacher's attention focus (on achievement) would typically prevent her from effectively assessing all the other variables at the same time. Third, even when multivariate information is available and the teacher is able to jointly attend to and encode the multivariate contingency data, memory restrictions would prevent her from remembering the full multi-dimensional distribution. Fourth, the time and patience needed to fill such a monstrous array with data would paralize the teacher. Before the rarest cells of the giant design are filled with observations, the school year would be over.

Thus, closer analysis of the information input from which correlations have to be inferred reveals that the notion of multivariate observations, which is so familiar from statistics courses, may be far away from the real empirical world. Exactly because extensional information about joint frequencies is often not available, several authors have emphasized that causal inferences and contingency judgments have to rely on intensional information and spatial-temporal contiguity (Chater & Oaksford, 2006; Fiedler, 2000a).

*PC inferences afford a viable alternative*. However, the PC illusion suggests an alternative. Even when joint frequencies of the full Cartesian product of all event combinations cannot be used, statistical information may still be used at a realistic level. Just like animals can naturally learn statistical proportions, such as the reinforcement rate associated with specific ecologies, or the baserate of a signal or conditional stimulus, the teacher can be assumed to be quite effective at learning the baserates of observations for many attributes of interest: a student's baserate of correct responses as an indicator of achievement; the student's rate of raising hands as an indicator of motivation; the relative number of TV-related remarks from that student etc. To be sure, these learned proportions or baserates cannot be interpreted on an absolute scale; they only provide ordinal information about the relative prevalence of an attribute across different ecologies. A particular student is characterized by a high baserate of correct responses (high achievement) and, observed on a different occasion, a high baserate of raising hands (motivation), or by a relatively low proportion of TV comments (low TV consumption). There is no evidence that individual responses are more correct when the same student raised her hand or on particular days following high TV consumption. Rather, at a higher level of aggregation, the teacher combines a high achievement baserate with a high motivation baserate, or a low TV baserate, in a PC-type inference. Similarly, at the level of classes, when the achievement baserate is high and the motivation baserate is also high, the co-occurrence of two baserates is used for inferring a positive contingency between motivation and achievement.

*PC inferences in empirical research*. From this sketch, it is but a small step to realizing that our teacher basically applies the same rationale that empirical scientists use in a probabilistic world that calls for analyses of group data to filter out noise and unreliability. For example, researchers compare student groups with different baserates of TV consumption and conclude from differential aggression baserates in both groups that TV enhances aggression. However, the PC inference is not only common in correlational research but even extends to empirical research. In a typical experimental setting, a manipulation (e.g., films to induce good mood) is administered to an experimental group but not to a control group. A manipulation check is then used to ensure that a majority of participants in the experimental group shows the intended change on an independent variable (mood). If a majority of participants in the experimental group exhibits an effect in the dependent variable (e.g., increased top-down inferences; Fiedler, 2001), researchers assume to have demonstrated an individual-level causal influence of mood on cognition: positive mood causes top-down thinking. However, the ecological correlation between mood and cognition baserates at group level does not logically imply that mood and cognition are related within individuals (e.g., watching the film may have caused good mood in most participants, and the same film may have also caused a procedural priming of top-down thinking, but independently of the mood effect). Whatever the real impact of the group treatment was, it may have affected mood and cognition independently, inducing the same tendency toward top-down thinking in both subsets of good-mood and bad-mood participants. To repeat, two baserates do not make up a contingency proper.

PC-like thinking is also common when scientists engage in theoretical reasoning. Theoretical models often involve more variables than can be controlled in singular experiments. Facing this situation, empirical tests of a theory linking a dependent variable Y to, say, four independent variables U,V,W,X, are based on different experiments, each of which includes a different subset of, say, two independent variables: U,V; U,W; U,X and so forth. Researchers then combine the findings obtained in two-factorial designs to inferences about a four-factorial theory. Logically, this inference from two-factorial relations to a four-factorial relation reflects the same category mistake as the elementary PC inference from two main effects (baserate tendencies) to an interaction (contingency).

Thus, when everyday judges and decision makers fall prey to PC effects, they seem to adapt to hard constraints of the information ecology, which also force scientists to resort to the same inference scheme. Just because in reality complete multivariate data arrays are either not available, or because the assessment and memorization of such highly demanding data is not feasible, a plausible and economical heuristic is to resort to sizeable contingencies at aggregate level. Rather than assessing the contingency of student achievement and TV consumption (in addition to SES, motivation, parent's profession etc.) over individual students, the teacher analyzes the relation between baserates of all these variables in different ecologies (e.g., groups, categories, time aggregates). Likewise, researchers take the correlation of group baserates as evidence for intra-individual processes. All this may be considered a normal manifestation of bounded rationality (Simon, 1956) – a rationality that is bounded by the accessibility constraints of the environment and by the cognitive-capacity constraints of the human mind.

*Functional value of PCs*. As already pointed out, there is often no alternative to using aggregate data as a proxy for individuating processes, that is, to making inferences from categories to individual cases. However, the crucial question is whether the PC proxy is functional, informing correct predictions and decisions. The above allusion to the analogy between PCs and scientific inference already suggests that the proxy cannot be that irrational. Prudent theorists (e.g., Huttenlocher, Hedges & Vevea, 2000) have made a strong point arguing that reliance on category knowledge may often inform rational inferences. However, explicating the functionality of PC illusions is not that easy. The utility (i.e., the benefits and costs) of judges' reliance on group baserates or averages depends on several considerations.

First of all, there is no a priori ground to assume that any particular aggregation level is the ultimately true or most useful level. PC illusions shift attention toward contingencies that hold at higher rather than at lower levels of aggregation. But what can be said about the functionality of such a shift? – One asset, already pointed out, is that higher levels of aggregation make assessment possible at all. Another obvious asset of aggregation is to increase the reliability of observations in a fallible, noisy world. Sill another, related advantage is that regularities observed at a higher level are more general and less restricted to the peculiarities of higher-order interactions and particular cases. However, aside from these apparent assets, it is worth while speculating about the systematic influence that a bias toward higher aggregate levels can have in the long run.

Crucial to adaptive cognition is the prediction and control of the origins of positive and negative payoffs. A cognitive module that supports the formation of higher aggregates (e.g., averages over longer time segments) forces the organism to attend more to long-term, global payoffs than to short-term, local payoffs. This can be of considerable value in overcoming delay-of-gratification problems, which is a precondition for long-term adaptive behavior.

In a similar vein, causal influences may be induced more effectively and interventions may often be more feasible at the aggregate level of ecologies than at the level of particular individuals. Just as experimental treatments do not warrant the same influence on every individual but only an average influence on a randomized group, many everyday interventions may be more easily applied to ecologies than to individuals. The teacher can make her lesson more interesting or change her teaching style for the whole class. Purchasing a TV set changes the ecology rather than a specific student. Similarly, for an animal to survive, it is typically more feasible to avoid certain ecologies than to try to change an individual predator. Or, for a consumer to reduce the consumption costs, she should search for a less expensive market rather than trying to negotiate the price of individual products. Anyway, to arrive at an informed analysis of the adaptive value of PC illusions, one has to engage in a systematic analysis of the payoff structure of the environment – which is a demanding theoretical task.

As usual, the various benefits of aggregate-level assessment come along with distinct costs. To the extent that ecological class differences intrude into the teacher's evaluation of individual students' performance, of course, evaluation becomes unfair and biased. After all, what has to be evaluated is individual students' performance, independent of the class. In this regard, it cannot be denied that PCs, like all cognitive illusions, turn out to produce erroneous results when carried over to new task settings, for which they are not functional. Nevertheless, from a more distant perspective, interpreting individual students' achievement in the context of their class environment is not that irrational, for the causes of achievement may be found at class level (teacher, group behavior, subject matter) and appropriate interventions may also lie at class level. What is good and effective from a systemic or evolutionary point of view may not always appear "fair" or "just" at individual level.

*Failures to Aggregate: Factors that Counteract PC Illusions*

If an analysis of the learning context of contingency assessment renders the PC illusion plausible, then by the same token the learning environment may also explain those conditions that counteract PC illusions. Recall that the overview of tricky paradigms at the outset included several phenomena that run opposite to the PC effect, reflecting failures to consider aggregate information. For instance, suboptimal choices in dilemma games originate in the failure to understand that cooperation is of great advantage across many trials, although defection is clearly the optimal strategy at the level of individual trials. A similar failure to aggregate over longer time frames is apparent in various delay-of-gratification problems (Metcalf & Mischel, 1999). In group decision making, too, performance suffers from the fact that the information that is distributed over group members is not combined effectively (Mojzisch & Schulz-Hardt, 2006). Thus, the PC bias toward higher aggregation levels is not universal but restricted to certain task conditions. But what are the boundary conditions that trigger either PC-like biases toward higher aggregation levels or reverse biases toward low-aggregate, individuating information in different task settings?

*Encoding and reinforcement structure of the task.* There is little direct empirical evidence at the moment to provide an informed answer, but two crucial boundary conditions suggest themselves, the encoding structure and the reinforcement structure of the task. For a general rule, PC illusions can be expected to occur under conditions that facilitate aggregate-level encoding and aggregate-level reinforcement. In contrast, when the task environment emphasizes individual events or outcomes and prevents the decision maker from aggregate-level encoding and reinforcement, then an opposite bias can be expected to occur.

To illustrate this crucial point, consider the hypothetical dilemma game depicted in Figure 4. Given nature plays cooperatively, the participant wins 10 from cooperation but 20 from defecting on every trial. Given nature defects, the participant wins nothing (0) from cooperation but 2 from defecting. Thus, at trial level, one ought to defect. However, it is well known that multi-trial dilemmas create reciprocal behavior, that is, on aggregate, across trials, nature tends to match one's own strategy, playing tit for tat. Let us assume nature cooperates 75% of the time when participant also cooperates on 75% of the trials, and that nature's cooperation rate is 25% when the participant's cooperation rate is 25%. As Figure 4 shows, when aggregating over many trials, the expected payoff is clearly higher (i.e., 5.625 vs. 1.874) when the cooperation baserate is high (75%) rather than low (25%). The question then is whether participants have a real chance to encode this aggregate-level contingency and to vividly experience the reinforcement associated with aggregate-level strategies.

In a typical dilemma game environment, participants have to make a decision on every single trial, and they are immediately reinforced with a feedback about the outcome of that trial. Such immediate reinforcement forces the participant not to forgo any profit at trial level and prevents her from costly long-term explorations at aggregate level. In order to experience and encode the aggregate-level contingency in favor of cooperation, it would be necessary that (a) the participant refrains from maximal payoffs over an extended time period; (b) that nature reciprocates and also converts to cooperation; (c) that the participant must somehow anticipate reciprocal cooperation; and (d) that learning and memory of the contingency between one's own cooperation baserate and nature's reciprocation baserate has to be successful. No doubt, these conditions are very unlikely to be met simultaneously.

Conversely, a slightly modified version of the very same dilemma task may indeed produce a PC-effect, facilitating the insight that cooperation is worth while. Let us assume the participant is not an actor in a dilemma task who is reinforced on every trial but, rather, an observer who witnesses an actor’s cooperation rate and reward rate over a longer time period, postponing a judgment to the end of the entire sequence. From such a remote perspective, the participant should easily recognize that a high baserate of cooperation comes along with a high payoff rate, especially when other strategies (e.g., high defection baserates observed in other players or in different time periods) are met with low payoffs or losses.

To continue this thought experiment, moreover, closer analyses may reveal that the wise observer uses the PC algorithm rather than a contingency algorithm proper. That is, it may be sufficient to recognize that both cooperation baserates, for the player and for nature, are skewed in the same direction, regardless of whether the player’s and nature’s cooperation actually correlate over trials. To test this assumption one might let observers witness a sequential dilemma game in which both the player and nature cooperate at a high (75%) or both at a low (25%) baserate. However, in one experimental condition, nature cooperates clearly more when the player cooperates. In another condition, nature cooperates at the same (constantly high or low) rate regardless of whether the player cooperates or not. If the observer’s belief that payoff increases with cooperation is the same in both conditions, this would be cogent evidence for PC inferences rather than contingencies proper.

Further analyses of other multi-level problems corroborate the assumption that the spontaneously chosen aggregation level reflects the encoding and reinforcement structure of the task. In group decision making, what is most likely to be encoded and communicated in group discussion is the individual decision maker’s personal preferences. In contrast, the group-level information and the group-level preference is unlikely to be encoded, discussed and assessed effectively (Mojzisch & Schulz-Hardt, 2006). With respect to reinforcement, or payoffs, although the modal individual preferences may diverge from the aggregate group preferences, decisions based on simple majority rules, such as the Condorcet principle[5], have been shown to provide close to optimal solutions most of the time (Hastie & Kameda, 2005). Moreover, the unreflected premise in group-decision making research that group-level information is more valid than a majority rule applied to individual-level information is an open empirical question, rather than an *a-priori* truth. In any case, it is no surprise that group decisions do not exhibit PC illusions (i.e., no bias toward aggregate-level information), simply because the encoding structure of the task setting does not support aggregate-level encoding.

The crucial point to be conveyed here – that the encoding and payoff structure of the task determines the aggregation level of the decision process – is nicely illustrated in recent research on Simpson's paradox. In a typical experiment (Fiedler, Walther, Freytag & Nickel, 2003; Schaller, 1992; Waldmann & Hagmayer, 1995), participants observe, as already depicted above, that more female applicants for graduate programs than male applicants are rejected. However, the seeming disadvantage of females turns out to reflect an ecological correlation; that is, what renders females less successful is the higher rejection rate of those universities to which females apply predominantly. When the unequal rejection rate of different universities (ecologies) is partialled out, the rejection rate of female individuals within universities actually turns out to be lower than the male rejection rate. Thus, Simpson's paradox is a special case of a multi-level problem that entails a spurious correlation.

Thus, participants who have to judge and compare males and females on such complex tasks have to make a choice between two representations: (a) they can either encode the female disadvantage across universities (noting that rejection rate in some universities are high because there are to many female applicants), or (b) encode the female advantage within universities (noting that the apparent male superiority merely reflects the unequal rejection rates of different universities). What level is rational, or normatively correct, depends on one's causal model. If the cause lies in the universities' unequal difficulty level, it is rational

[5] According to the Condorcet rule, a choice option or candidate is chosen if it receives more than half of the individual votes in group decision making.

to compare individual males and females within universities. If the cause lies in the universities' unequal gender composition, it is rational to focus on the ecological correlation between gender rates and rejection rates across universities.

What cognitive representation is chosen in judgment experiments using Simpson's paradox depends on the encoding and reinforcement structure. On one hand, with regard to encoding, when the temporal order in which the stimulus observations are provided supports interpretation (a), presenting applicant's gender before the university name to highlight the primacy of applicant gender as an antecedent of university rejection rates, then they tend to see a female disadvantage. If, however, the presentation order facilitates interpretation (b), indicating the university prior to the applicant's gender to highlight the antecedent role of university standards, they tend to recognize the female superiority (Fiedler et al., 2003; Fiedler, Walther, Freytag & Stryczek, 2002).

On the other hand, with regard to reinforcement, when feminist motives are solicited or when participants hold feminist attitudes, they tend to prefer interpretation (b) over (a) because the former interpretation is more reinforcing from a feminist perspective (Schaller, 1992). – Needless to repeat that the "true" solution to the problem is unknown, or at least it cannot be determined on the basis of statistical contingencies alone.

*Higher-order memory codes*. PC-like biases toward higher aggregation levels are most pronounced when complex information calls for higher-order categorical encoding. Participants in a study by Fiedler and Graf (1990) first learned whether a virus was observed or not in 24 different countries. Then, in a second run, they were informed about the occurrence of a disease in the same countries. To be sure, memorizing the precise distribution of virus and disease across as many as 24 countries is hardly possible. However, given that all countries could be categorized into six geographical clusters (Scandinavian; Mediterranean; South American etc.), the memory load reduced to learning the relative occurrence rate of virus and disease in only six geographical clusters. These were spontaneously used as highly effective and economical encoding units, as often demonstrated in memory experiments with categorized lists (Cohen, 1969; Shuell, 1969). Thus, holding the contingency between virus and disease constant (i.e., the number of matching pairs, or countries in which virus and disease were jointly present or absent), participants came up with pronounced contingency estimates only when matching pairs consistently came from the same clusters (i.e., virus and disease jointly present in all countries of some clusters and jointly absent in others), but not when the same number of matches was evenly distributed across all clusters. In the former case, a marked ecological correlation helped participants to encode the coincidence of virus and disease at the level of higher-order memory units.

*Summary and Conclusions*

The present chapter started with the contention that the complex and difficult problems of the environment that the probabilistic mind has to deal with do not primarily reflect the probabilistic nature of the empirical reality. That the world is not strictly deterministic not only creates uncertainty and sometimes stress but also entails optimism and the potential for progress and control. Rather, what renders the world difficult and conflict-prone is that it looks different from different perspectives. This important insight is at the heart of several research paradigms that have enhanced our understanding of the probabilistic mind. The conditional reasoning paradigm highlights the fact that inferences from X to Y may diverge drastically from reverse inferences from Y to X (Fiedler, 2000b; Koriat, Fiedler & Bjork, 2006). Construal-level theory is concerned with the changing appearance of the world as a function of temporal, spatial and social distance (Trope & Liberman, 2003). The pseudocontingency (PC) illusion that was the focus of the present chapter adds another way in which the world is subject to perspectival changes and relativity (see also Stewart, Chater, Stott & Reimers, 2003). As a matter of principle, environmental correlations vary in size and even in sign when considered at different aggregation levels. The PC illusion reflects a cognitive bias toward assessing contingencies at high rather than low levels of aggregation.

The contingencies that hold between group or category baserates are (mis)taken as a proxy for the contingency that holds between individuating people or events within categories.

The PC illusion was only recently discovered, but it was then found to generalize across many task conditions, content areas, and decision problems. Like all illusions, the impact of PC biases can be quite massive and hard to believe. However, just like other perceptual and cognitive illusions, PCs can be understood as plausible and functional when their learning environment is taken into account. Like any illusion, PCs can be characterized as overgeneralizations of heuristics that function well in many task contexts while producing errors and distortions when carried over to other contexts.

In any case, PCs constitute a fascinating topic in the study of the probabilistic mind, the topic of higher-order contingency problems. A closer examination of other examples of such higher-order contingencies – such as Simpson's paradox, dilemma games, or group decision making – suggests that PC biases to attend to high aggregation levels may be reduced, eliminated or even reversed when the encoding and reinforcement structure of the task facilitates an attention shift from high levels to low levels of aggregation.

Research on higher-order contingency problems is only beginning to grow (Fiedler & Plessner, in press; Spellman, 1996). However, in spite of the paucity of systematic research conducted so far, there can be no doubt that such problems provide a major challenge for the probabilistic mind as it has to cope with the pitfalls of utility assessment, risky choice, causal inference and prediction and – last but not least – with the pitfalls of scientific inference.

References

Allan, L.G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin*, *114*, 435-448.

Alloy, L.B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: the joint influence of prior expectations and current situational information. *Psychological Review, 91*, 112-149.

Blümke, M., & Fiedler, K. (2007). *Base rate effects in the IAT*. Manuscript submitted for publication.

Chater, N. & Oaksford, M. (2006). Mental mechanisms: Speculations on human causal learning and rasoning. In K. Fiedler & P. Juslin (Eds.), *Sampling and adaptive cognition* (pp. 210-236). New York: Cambridge University Press.

Chohen, B.H. (1966). Some-or-none characteristics pf coding behavior. *Journal of Verbal Learning and Verbal Behavior, 5*, 182-187.

Fiedler, K. (2000a). Illusory correlations: A simple associative algorithm provides a convergent account of seemingly divergent paradigms. *Review of General Psychology, 4*, 25-58.

Fiedler, K. (2000b). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review, 107*, 659-676.

Fiedler, K. (2001). Affective states trigger processes of assimilation and accommodation. In L.L. Martin & G.L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook.* (pp. 85-98). Mahwah, NJ: Lawrence Erlbaum Associates.

Fiedler, K., Blümke, M., & Unkelbach, C. (2007). *Impact of valence baserates on evaluative priming*. Unpublished research, University of Heidelberg.

Fiedler, K., & Freytag, P. (2003). Social judgments based on pseudocontingencies: A forgotten phenomenon. In J.P. Forgas, K.D. Williams, & W. von Hippel (Eds.), *Social judgments: Implicit and explicit processes* (pp. 162-179). New York: Cambridge University

Press.

Fiedler, K., & Freytag, P. (2004). Pseudocontingencies. *Journal of Personality and Social Psychology*, 87, 453-467.

Fiedler, K., Freytag, P., Randoll, A., & Vogel, T. (2007). *Pseudocontingencies in the daignostic assessment of questionnaire responses*. Unpublished research, University of Heidelberg.

Fiedler, K., Freytag, P., & Unkelbach, C. (2007). Pseudocontingencies in a simulated classroom. *Journal of Personality and Social Psychology, 92*, 665-667.

Fiedler, K., & Graf, R. (1990). Grouping and categorization in judgments of contingency. In J.-P. Caverni & J.-M. Fabre (eds), *Cognitive biases* (pp. 47-57). Oxford: North-Holland.

Fiedler, K., & Plessner, H. (in press). Induction: From Simple Categorization to Higher-Order Inference Problems. In F. Strack & J. Förster (Eds.), *Social Cognition – the Basis of Human Interaction.* New York: Psychology Press.

Fiedler, K., Walther, E., Freytag, P., & Nickel, S. (2003). Inductive reasoning and judgment interference: Experiments on Simpson's paradox. *Personality and Social Psychology Bulletin, 29,* 14-27.

Fiedler, K., Walther, E., Freytag, P., & Stryczek, E. (2002). Playing mating games in foreign cultures: A conceptual framework and an experimental paradigm for trivariate statistical inference. *Journal of Experimental Social Psychology, 38,* 14-30.

Freytag, P. (2003). *Contextually determined typicality*. Berlin: Logos.

Goodie, A.S., & Fantino, E. (1996). Learning to commit or avoid the base-rate error. *Nature, 380*, 247-249.

Hammond, J.L. (1973). Two sources of error in ecological correlations. *American Sociological Review, 38*, 764-777.

Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. Psychological Review, 112, 494-508.

Haselton, M.G., & Funder, D.C. (in press). The evolution of accuracy and bias in social judgment. In M. Schaller, J.A. Simpson, & D.T. Kenrick (Eds.), *Evolution and social psychology*. New York: Psychology Press.

Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 566-581.

Huttenlocher, J. & Hedges, L., & Vevea, J.L. (2000). Why Do Categories Affect Stimulus Judgement? *Journal of Experimental Psychology: General, 129,* 220-241.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*, 211-228.

Koriat, A., Fiedler, K., & Bjork, R.A. (2006). Inflation of conditional predictions. *Journal of Experimental Psychology: General, 135*, 429-447.

Kutzner, F., Freytag, P., Vogel, T., & Fiedler, K. (2007). *Base-rate neglect based on base-rates in experience-based contingency learning*. Manuscript submitted for publication.

McKenzie, C. R. M. (1994). The accuracy of intuitive judgment strategies: Covariation assessment and Bayesian inference. *Cognitive Psychology, 26*, 209-239.

Meiser, T. (2006). Contingency learning and biased group impressions. In K. Fiedler & P. Juslin (Eds.), *Sampling and adaptive cognition* (pp. 183-209). New York: Cambridge University Press.

Meiser, T., & Hewstone, M. (2004). Cognitive processes in stereotype formation: The role of correct contingency learning for biased group judgments. *Journal of Personality and Social Psychology, 87*, 599-614.

Metcalfe, J., & Mischel, W. (1999). A hot-cool analysis of the delay of gratification: the dynamics of willpower. *Psychological Reiew, 106*, 3-19.

Mojzisch, A., & Schulz-Hardt, S. (2006). Information sampling in group decision making. In K. Fiedler & P. Juslin (Eds.), *Sampling and adaptive cognition* (pp. 299-326). New York: Cambridge University Press.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608-631.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, *15*, 351–357.

Schaller, M. (1992). In-group favoritism and statistical reasoning in social inference: Implications for formation and maintenance of group stereotypes. *Journal of Personality and Social Psychology, 63*, 61-74.

Schelling, T.C. (1978). *Micromotives and macrobehavior*. New York: Norton.

Shuell, T.J. (1969). Clustering and organization in free recall. *Psychological Bulletin, 72*, 353-374.

Simon, H.A. (1956). Rational choice and the structure of environments. *Psychological Review, 63*, 129-138.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Ser, B, 13,* 238-241.

Spellman, B.A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7,* 337-342.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling in group discussion. *Journal of Personality and Social Psychology, 48*, 1467-1478.

Stewart, N., Chater, N., Stott, H.P., & Reimers, S. (2003). Psospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General, 132*, 23-46.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110*, 403-421.

Waldmann, M., & Hagmayer (1995). Causal paradox: When a cause simultaneously produces and prevent an effect. *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

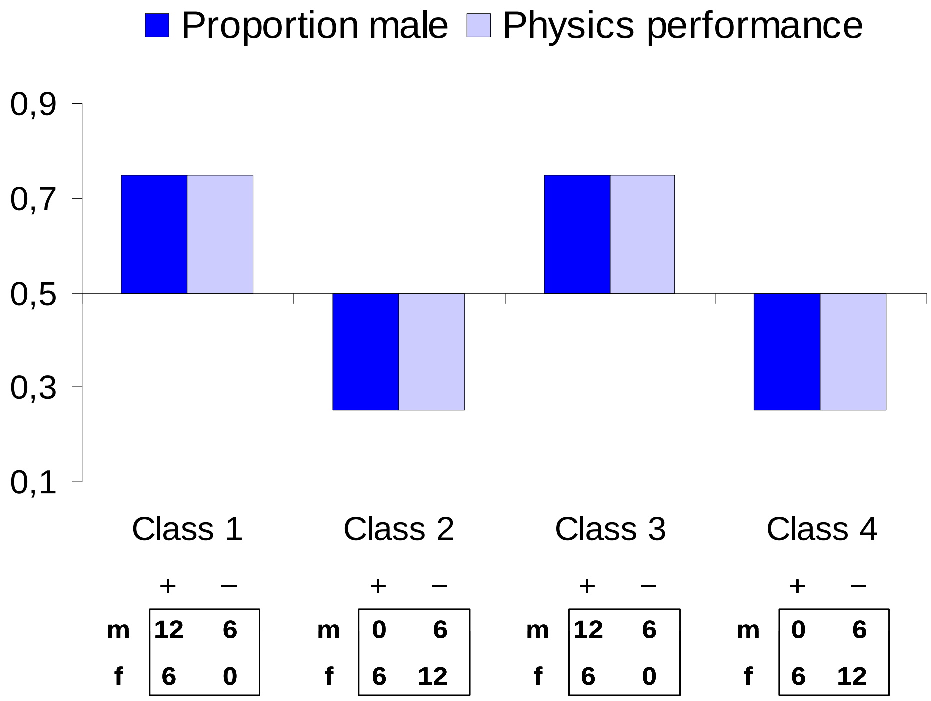


*Figure 1*

Illustration of divergent contingencies between student gender (m = male, f = female) and achievement (+ = high, – = low). At the level of classes, average performance is perfectly correlated with male proportion. However, within all classes, females outperform males.

| | $a+c=$ 100 | $b+d=$ 50 | Marginal baserate distributions | Contingency |
|---|---|---|---|---|
| I | $a=60$ | $b=40$ | $100 = a+b$ | $r = -.20$ |
| | $c=40$ | $d=10$ | $50 = c+d$ | |
| II | $a=67$ | $b=33$ | $100 = a+b$ | $r \sim .00$ |
| | $c=33$ | $d=17$ | $50 = c+d$ | |
| III | $a=75$ | $b=25$ | $100 = a+b$ | $r = +.25$ |
| | $c=25$ | $d=25$ | $50 = c+d$ | |

*Figure 2*

Setting pseudocontingencies apart from genuine contingencies. All three distributions imply positive pseudocontingencies, because the marginal distributions for rows and columns are skewed in the same direction, the contingency varies from negative (a) to zero (b) to positive (c).

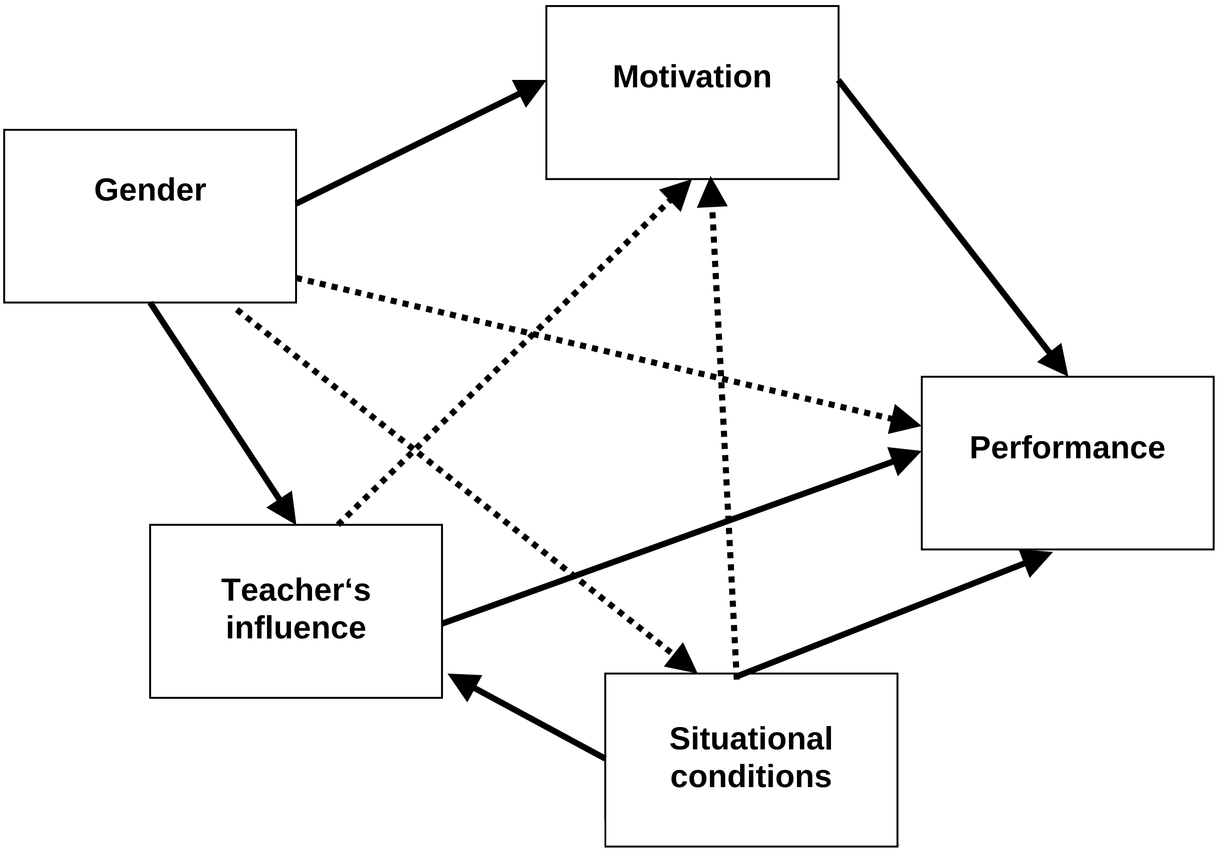


*Figure 3*

Variety of factors bearing contingencies with student performance

| | Nature cooperates | Nature defects |
|---|---|---|
| Player cooperates | 10 | 0 |
| Player Defects | 20 | 2 |

**Payoff matrix at the level of individual trials**

| | |
|---|---|
| Player cooperates | 5.625 |
| Player Defects | 1.874 |

**Expected value in a "friendly" environment. Nature cooperates 75% of the time. 25% defection**

*Figure 4*

Dilemma game as a multi-level problem. Although payoffs at the level of individual trials are higher when the player defects (upper panel), the aggregate value of cooperation is higher over many trials.

| Nr. | Author | Title |
|---|---|---|
| 07-54 | Klaus Fiedler | Pseudocontingencies - A key paradigm for understanding adaptive cognition |
| 07-53 | Florian Kutzner<br>Peter Freytag<br>Tobias Vogel<br>Klaus Fiedler | Base-rate neglect based on base-rates in experience-based contingency learning |
| 07-52 | Klaus Fiedler<br>Yaakov Kareev | Implications and Ramifications of a Sample-Size Approach to Intuition |
| 07-51 | Klaus Fiedler | The Ultimate Sampling Dilemma in Experience-Based Decision Making |
| 07-50 | Jürgen Eichberger<br>David Kelsey | Ambiguity |
| 07-49 | Tri Vi Dang | Information Acquisition in Double Auctions |
| 07-48 | Clemens Kroneberg | Wertrationalität und das Modell der Frame-Selektion |
| 07-47 | Dirk Simons<br>Nicole Zein | Audit market segmentation and audit quality |
| 07-46 | Sina Borgsen<br>Martin Weber | False Consensus and the Role of Ambiguity in Predictions of Othersí Risky Preferences |
| 07-45 | Martin Weber<br>Frank Welfens | An Individual Level Analysis of the Disposition Effect: Empirical and Experimental Evidence |
| 07-44 | Martin Weber<br>Frank Welfens | The Repurchase Behavior of Individual Investors: An Experimental Investigation |
| 07-43 | Manel Baucells<br>Martin Weber<br>Frank Welfens | Reference Point Formation Over Time: A Weighting Function Approach |
| 07-42 | Martin Weber<br>Frank Welfens | How do Markets React to Fundamental Shocks? An Experimental Analysis on Underreaction and Momentum |

| Nr. | Author | Title |
|---|---|---|
| 07-41 | Ernst Maug<br>Ingolf Dittmann | Lower Salaries and No Options: The Optimal Structure of Executive Pay |
| 07-40 | Ernst Maug<br>Ingolf Dittmann<br>Christoph Schneider | Bankers and the Performance of German Firms |
| 07-39 | Michael Ebert<br>Nicole Zein | Wertorientierte Vergütung des Aufsichtsrats - Auswirkungen auf den Unternehmenswert |
| 07-38 | Ingolf Dittmann<br>Ernst Maug<br>Christoph Schneider | How Preussag became TUI: Kissing too Many Toads Can Make You a Toad |
| 07-37 | Ingolf Dittmann<br>Ernst Maug | Valuation Biases, Error Measures, and the Conglomerate Discount |
| 07-36 | Ingolf Dittmann<br>Ernst Maug<br>Oliver Spalt | Executive Stock Options when Managers are Loss-Averse |
| 07-35 | Ernst Maug<br>Kristian Rydqvist | Do Shareholders Vote Strategically? Voting Behavior, Proposal Screening, and Majority Rules |
| 07-34 | Ernst Maug<br>Abraham Ackerman | Insider Trading Legislation and Acquisition Announcements: Do Laws Matter? |
| 07-33 | Dirk Simons | Independence, low balling and learning effects |
| 07-32 | Rainer Greifeneder<br>Herbert Bless | Relying on accessible content versus accessibility experiences: The case of processing capacity |
| 07-31 | Rainer Greifeneder<br>Herbert Bless | Depression and reliance on ease-of-retrieval experiences |
| 07-30 | Florian Heiss<br>Axel Börsch-Supan<br>Michael Hurd<br>David Wise | Pathways to Disability: Predicting Health Trajectories |
| 07-29 | Axel Börsch-Supan<br>Alexander Ludwig<br>Mathias Sommer | Aging and Asset Prices |

| Nr. | Author | Title |
|---|---|---|
| 07-28 | Axel Börsch-Supan | GLOBAL AGING - Issues, Answers, More Questions |
| 07-27 | Axel Börsch-Supan | MIND THE GAP: THE EFFECTIVENESS OF INCENTIVES TO BOOST RETIREMENT SAVING IN EUROPE |
| 07-26 | Axel Börsch-Supan | Labor market effects of population aging |
| 07-25 | Axel Börsch-Supan | Rational Pension Reform |
| 07-24 | Axel Börsch-Supan | European welfare state regimes and their generosity towards the elderly |
| 07-23 | Axel Börsch-Supan | Work Disability, Health, and Incentive Effects |
| 07-22 | Tobias Greitemeyer<br>Rainer Greifeneder | Why the Euro looked like a price booster: Differential perception of increasing versus decreasing prices |
| 07-21 | Patrick A. Müller<br>Rainer Greifeneder<br>Dagmar Stahlberg<br>Herbert Bless | Relying on accessibility experiences in procedural fairness judgments |
| 07-20 | Volker Stocké | The Motive for Status Maintenance and Inequality in Educational Decisions. Which of the Parents Defines the Reference Point? |
| 07-19 | Jürgen Eichberger<br>David Kelsey<br>Burkhard Schipper | Ambiguity and Social Interaction |
| 07-18 | Jürgen Eichberger<br>Willy Spanjers | Liquidity and Ambiguity: Banks or Asset Markets? |
| 07-17 | Patrick A. Müller<br>Jana Janßen<br>Dominique Jarzina | Applicantsí reactions to selection procedures ñ Prediction uncertainty as a moderator of the relationship between procedural fairness and organizational attractiveness |

| Nr. | Author | Title |
|---|---|---|
| 07-16 | Patrick A. Müller<br>Dagmar Stahlberg | The Role of Surprise in Hindsight Bias ñ A Metacognitive Model of Reduced and Reversed Hindsight Bias |
| 07-15 | Axel Börsch-Supan<br>Anette Reil-Held<br>Daniel Schunk | Das Sparverhalten deutscher Haushalte: Erste Erfahrungen mit der Riester-Rente |
| 07-14 | Axel Börsch-Supan<br>Dirk Krüger<br>Alexander Ludwig | Demographic Change, Relative Factor Prices, International Capital Flows, and their Differential Effects on the Welfare of Generations |
| 07-13 | Melanie Lührmann | Consumer Expenditures and Home Production at Retirement: New Evidence from Germany |
| 07-12 | Axel Börsch-Supan<br>Anette Reil-Held<br>Christina Wilke | Zur Sozialversicherungsfreiheit der Entgeltumwandlung |
| 07-11 | Alexander Ludwig<br>Dirk Krüger | On the Consequences of Demographic Change for Rates of Returns to Capital, and the Distribution of Wealth and Welfare |
| 07-10 | Daniel Schunk | What Determines the Saving Behavior of German Households? An Examination of Saving Motives and Saving Decisions |
| 07-09 | Axel Börsch-Supan<br>Anette Reil-Held<br>Christina Wilke | How an Unfunded Pension System looks like Defined Benefits but works like Defined Contributions: The German Pension Reform |
| 07-08 | Daniel Schunk | The German SAVE survey: documentation and methodology |
| 07-07 | Hans-Martin von Gaudecker<br>Carsten Weber | Mandatory unisex policies and annuity pricing: quasi-experimental evidence from Germany |
| 07-06 | Daniel Schunk | A Markov Chain Monte Carlo Multiple Imputation Procedure for Dealing with Item Nonresponse in the German SAVE Survey |
| 07-05 | Hans-Martin von Gaudecker<br>Rembrandt Scholz | Lifetime Earnings and Life Expectancy |

| Nr. | Author | Title |
|---|---|---|
| 07-04 | Christopher Koch<br>Daniel Schunk | The Case for Limited Auditor Liability - The Effects of Liability Size on Risk Aversion and Ambiguity Aversion |
| 07-03 | Siegfried K. Berninghaus<br>Werner Gueth<br>M. Vittoria Levati<br>Jianying Qiu | Satisficing in sales competition: experimental evidence |
| 07-02 | Jannis Bischof<br>Michael Ebert | Inconsistent measurement and disclosure of non-contingent financial derivatives under IFRS: A behavioral perspective |
| 07-01 | Jörg Oechssler<br>Carsten Schmidt<br>Wendelin Schnedler | Asset Bubbles without Dividends - An Experiment |
| 06-16 | Siegfried K. Berninghaus<br>Hans Haller | Pairwise Interaction on Random Graphs |
| 06-15 | Markus Glaser<br>Philipp Schmitz | Privatanleger am Optionsscheinmarkt |
| 06-14 | Daniel Houser<br>Daniel Schunk<br>Joachim Winter | Trust Games Measure Trust |
| 06-13 | Markus Glaser<br>Sebastian Müller | Der Diversification Discount in Deutschland: Existiert ein Bewertungsabschlag für diversifizierte Unternehmen? |
| 06-12 | Philipp Schmitz<br>Markus Glaser<br>Martin Weber | Individual Investor Sentiment and Stock Returns - What Do We Learn from Warrant Traders? |
| 06-11 | Siegfried K. Berninghaus<br>Sven Fischer<br>Werner Gueth | Do Social Networks Inspire Employment? - An Experimental Analysis - |
| 06-10 | Christopher Koch<br>Carsten Schmidt | Disclosing Conflict of Interest - Does Experience and Reputation Matter? |