SONDERFORSCHUNGSBEREICH 504

Rationalitätskonzepte,
Entscheidungsverhalten und
ökonomische Modellierung

No. 08-37

The reproduction of base-rates promotes pseudocontingencies

Peter Freytag*
and Tobias Vogel**
and Florian Kutzner***
and Klaus Fiedler****

December 2008

*Universität Heidelberg, email:

**Sonderforschungsbereich 504, email:

***Sonderforschungsbereich 504, email: florian.kutzner@psychologie.uni-heidelberg.de

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

Universität Mannheim
L 13,15
68131 Mannheim

Word count (text): 5963
Word count (abstract): 121

The reproduction of base-rates promotes pseudocontingencies

Peter Freytag Tobias Vogel

Florian Kutzner Klaus Fiedler

University of Heidelberg, Germany

Abstract

Fiedler and Freytag (2004) proposed an alternative pathway to contingency assessment in terms of pseudocontingencies (PCs). PCs reflect the utilization of base-rate information in the formation of contingency judgments. Here, we introduce an instantiation of the phenomenon based on the mere reproduction of the base rates. Using a relationship-counseling scenario, participants in two experiments produced positive correlations on both indirect and direct measures of the contingency between partners' responses to the subscales of a relationship inventory, although the objective contingency within each subscale had been negative in an initial learning phase. The magnitude of these effects was predicted accurately by computer simulations reproducing the base rate of 'yes' responses for each partner and domain. The findings are discussed within the PC framework.

The reproduction of base-rates promotes pseudocontingencies

Recently, Fiedler and Freytag (2003, 2004) proposed an alternative pathway to contingency assessment in terms of pseudocontingencies (PCs). PCs reflect the utilization of base-rate information in the formation of contingency judgments. Clear demonstrations of PCs arise when the actual contingency between two variables, as determined by cell frequency based algorithms (e.g., delta p, phi), has a sign opposite to the perceived contingency, as assessed by – or derived from – frequency estimates, cued recall tests, or conditional probability estimates.

In the case of bivariate frequency distributions, for instance, the perceived contingency of two dichotomous variables, X and Y (with levels X+, X-, Y+, and Y-), may be termed pseudocontingent if (a) the two base rates are skewed and (b) contingency judgments link the more frequent level of the one variable to the more frequent level of the other variable (e.g., Fiedler & Freytag, 2004; Fiedler, Freytag, & Unkelbach, 2007). In propositional form, the underlying judgment heuristic may read: "If the frequent (rare) level is observed in the one variable, then the frequent (rare) level is likely to be observed in other variable as well". The likelihood that people will engage in this process increases when the coordination of individual observations of X and Y is difficult (e.g., when X and Y values cannot be observed simultaneously) or even impossible (e.g., when the stimulus series reveals the base rates only).

Similarly, in the case of trivariate frequency distributions, strong PCs have been observed when the base rates of X and Y are skewed toward an elevated proportion of X+ and Y+ cases in context Z1, but toward an elevated proportion of X- and Y- cases in context Z2. Presumably, the variation of the base rates of both X and Y with the context variable Z enhances the salience of their being skewed. In line with this reasoning, the likelihood that people will form PCs in trivariate scenarios has been observed to increase with the accuracy with which participants learn the joint variation of the base rates of the focused variables with the context variable Z (cf. Meiser, 2007; Meiser & Hewstone, 2004).

Several examples of the findings typically obtained in experiments comprising one or two contexts can be found in a recent review by Fiedler, Freytag, and Meiser (in press). Here, we would like to introduce a novel instantiation by which the utilization of base-rate information can affect contingency perception, one that is based on the reproduction of base rates for as many as four contexts. Applying the PC framework to a new task situation requires making a clear distinction between (a) the pattern of base rates in the input information, (b) the cognitive processes transforming the pattern into some kind of response set, and (c) the resulting judgment biases. Note that it is the latter that we term a PC, because PCs extend to the case of a subjectively perceived contingency that reflects the utilization of base-rate information, irrespective of whether the objective contingency entailed in the stimulus information is of a sign opposite to the PC obtained or undefined (e.g., Fiedler & Freytag, 2004; Freytag & Fiedler, 2008). Importantly, the intervening cognitive process may be subject to variation with the characteristics of the task at hand (e.g., stimulus presentation formats, reinforcement schedules, or problem structure), with the only common denominator underlying all demonstrations of PCs being that perceived contingencies vary with the pattern of base rates.

The process we would like to propose here is best conveyed by a sketch of the scenario used in the experiments reported below: Imagine a psychologist specialized in partnership-counseling. In this domain, partners’ self-reports for dimensions of interest are routinely compared to each other, the rationale being that considerable disagreement may point to sources of conflict. Accordingly, our psychologist runs a relationship quality interview with each partner. Interestingly, both partners provided mostly ‘yes’ responses for some domains (e.g., joint activities and intimacy), and mostly ‘no’ responses for others (e.g., household and arguing).

The top panel of Figure 1 gives the joint frequencies of ‘yes’ and ‘no’ responses provided by the couple. In Domain 1, 6 items were endorsed by both partners, and 2 items each were either endorsed by her but denied by him, or denied by her but endorsed by him. Computing delta p with these cell frequencies, the within-domain contingency between the two partners’ responses is (6/8) - (2/2) = .75 - 1 = -.25. The same negative contingency also holds in each of the other domains of life. Note, however, that both partners endorsed most items in Domain 1 and in Domain 3, and that they denied most items in Domain 2 and in Domain 4. Aggregating responses across domains (i.e., summing up the cell entries obtained within the four domains), the contingency of the partners’ responses is thus (12/20) - (8/20) = .6 - .4 = .2. Taken together, there is a weak negative correlation at the domain level, a weak positive correlation at aggregate level, and there is a perfect positive ecological correlation between the base rates of ‘yes’ responses in the four domains.

As a practitioner, our psychologist is interested in the degree to which the two partners agree with each other in their assessment of the different domains of life. That is, she wants to know whether the probability of his endorsing items she endorsed exceeds the probability of his endorsing items she rejected. Viewed from a normative perspective, the appropriate basis for such a judgment is located at the item level. Given the complexity of the task (multiple items, multiple domains, multiple persons), however, she may be left with nothing but the summary profiles for his and her responses (i.e., with partners’ base rates of ‘yes’ responses in the four domains as encoded or retained in memory). Coordinating the numerous pieces of information would be bothersome, it may well be impossible if she tried to coordinate the individual answers from memory alone.

Yet, our psychologist cannot simply decline a reaction (e.g., cancel the entire therapy). What she can do, however, is to take a neutral position on conditional probability. When approaching contingency-related tasks, a seemingly rational strategy may consist in the reproduction of the base rate of ‘yes’ responses for each partner and domain, without any attempt to coordinate partners’ position on individual items. If both partners endorsed most intimacy items, for instance, she could assign a ‘yes’ to both the majority of her responses and the majority of his responses, selecting the items to be marked as endorsed by each partner at random. Under the conditions depicted in Figure 1, however, this seemingly neutral strategy is bound to produce pseudocontingent impressions. Given that the more prevalent response in each domain has a base rate of .8, the chances of obtaining the more frequent response for both partners in a given domain is .8 x .8 = .64, and the chances of obtaining the rare response for both target persons is .2 x .2 = .04. Collapsing data over domains, we get a probability of .68 that judgments generated using the base-rate reproduction strategy will fall into the a-d-diagonal of the aggregate level frequency table, thus inflating the proportion of observations supporting a positive contingency. At the domain level, moreover, the independent reproduction of the base rates is bound to produce zero contingencies (because the proportion of items endorsed by him will be the same for items she had endorsed as for items she had rejected), although in fact all domain-level contingencies had been negative.

To obtain a reliable estimate of the magnitude of perceived correlations based on base-rate reproduction, we ran a simulation programmed to apply the reproduction strategy to the first indirect measure used in Experiment 1, which asked for predictions of the two partners’ responses to four additional items in each domain. For 1000 simulated participants, the simulation randomly assigned the frequent answer in a given domain with a probability of .8, and the infrequent answer with a probability of .2, thus reproducing the base rates of frequent and infrequent answers.

If both partners had endorsed most items in a domain, that is, a random generator produced a ‘yes’ for her at a rate of .8 in response to each of the four items in a first run – as well as for him in a second run. And if both partners had rejected most items in a domain, a ‘yes’ was produced for her at a rate of .2 in a first run – and for him in a second run. Figure 2 gives the percentages of cases falling into different intervals of the aggregate delta p. The mean value of the distribution is .48, and the standard deviation is .18. The net effect of applying the seemingly neutral base-rate strategy, then, is that the aggregate correlation will be raised to a level of about .48 instead of .2, and that the within-domain correlation will be raised to a level of 0 instead of -.25.[1]

Experiment 1

The experiments reported below were designed to test the hypothesis that people adopt the base-rate reproduction strategy when confronted with jointly skewed base rates for two or more contexts, thereby producing correlated judgments at the aggregate level. Experiment 1 serves to introduce the experimental paradigm in more detail. The frequency distributions given in the top panel of Figure 1 were used for the stimulus series. Apart from the pseudocontingency-prone constellation in the stimulus material, we also manipulated two features of the experimental task intended to affect the degree to which participants would focus, and ultimately utilize, the information conveyed by the domain-wise base rates of ‘yes’ responses. First, the items pertaining to the different domains of life were either presented block-wise or in an alternating fashion. Second, written instructions focused participants’ attention to partners’ agreement at either the item level or the subscale level. As both block-wise presentation and subscale-level focus should facilitate the extraction of the base rates, their impact should be strongest in the block-wise/scale-focus condition. Finally, we counterbalanced the domains with elevated base rates of ‘yes’ responses.

That is, ‘yes’ base rates were high for joint activities and intimacy, but low for household and arguing for half the participants (good relationship condition), and vice versa for the other half of participants (bad relationship condition).

Method

Participants and design. A total of 64 undergraduate psychology students (32 women and 32 men; *M*(AGE) = 25.47, *SD*(AGE) = 6.82) were recruited for a study on clinical judgment in partial fulfillment of a course requirement. Under an equal *n* constraint, participants were randomly assigned to one of the eight experimental conditions resulting from the orthogonal variation of the between-participants variables presentation (alternate vs. block-wise), agreement (item-focus vs. scale-focus), and relationship quality (good vs. bad). Several contingency measures were administered within participants at each level of the within-participants variable domain (joint activities vs. household vs. intimacy vs. arguing).

Learning phase. Participants were recruited for a study on clinical judgment. Written instructions informed participants that higher levels of disagreement between self-assessment and assessment by others may point to potential sources of conflict in romantic relationships, and that the usefulness of clinical judgments of relationship quality depends on the accuracy with which counselors can extract the covariation of partners’ responses to different aspects of life. Participants learned that their first task was to study the answers provided by a real couple to a relationship inventory assessing four domains of life, with the explicit goal to monitor agreement either at the level of individual items (item-wise condition) or at the level of domains (scale-wise condition). Participants then selected one pair of envelopes from a set of eight. All envelopes were marked with a gender symbol and the alleged case number of the couple in an investigation run by the local center for psychotherapy.

Each subscale of the questionnaire comprised 10 items that could be endorsed or rejected, respectively, by checking either an ‘agree’ option or a ‘disagree’ option. In the block-wise condition, items were presented in one block per subscale, with a fixed random order (i.e., Domain 1: joint activities; Domain 2: household; Domain 3: intimacy; Domain 4: arguing). In the alternate condition, in contrast, each subset of four consecutive items featured one item per domain of life. In order to increase the credibility of the cover story, two different versions of the questionnaire were used, a self-assessment version filled in by the female partner and a partner-evaluation version filled in by the male partner. The only difference between the two versions consisted in the perspective of the items. For instance, the first intimacy item read, “I use to kiss him hello” in the self-assessment version, but “She uses to kiss me hello” in the partner-evaluation version.

The relationship quality variable determined whether or not endorsement base rates were high in domains with desirable aspects of life (i.e., joint activities and intimacy) and low in domains with undesirable aspects of life (i.e., household and arguing). Stimulus distributions were the same as in Figure 1 for participants in the good relationship condition. For participants in the bad relationship condition, the base rates of responding ‘yes’ were switched (i.e., high endorsement base rate in Domain 2 and in Domain 4, low endorsement base rate in Domain 1 and in Domain 3). Participants were instructed to study the female partner’s responses first, and to put her questionnaire back into its envelope before studying the male partner’s responses. Once participants had finished studying the questionnaires, they returned the envelopes to the experimenter, who provided them with a questionnaire containing the dependent measures.

Extrapolation task. The first questionnaire constituted an extrapolation task calling for an extrapolation of participants’ impression of the couple to additional items of the questionnaire. For each subscale of the inventory, the extrapolation task comprised four additional items, and participants were asked to fill in the answers they would expect the female partner and the male partner, respectively, to give on these items. For each item and for each partner, participants checked either the 'agree' option or the 'disagree' option. As in the learning phase, the female partner's form always preceded the male partner's form

Cued recall task. Upon completion of the extrapolation task, participants in all conditions were asked to reproduce the female partner's answers and the male partner's answers to the original questionnaires used in the learning phase. To this end, participants filled in a copy of the self-assessment version of the questionnaire for the female partner and a copy of the partner-evaluation version for the male partner. As before, responses for the female partner had to be completed before turning to the responses of the male partner.

Base-rate estimates. As a manipulation check, participants were asked to reproduce the base rates of endorsing responses separately for each combination of target person and domain. For instance, the item for her endorsement base rate in the intimacy domain read: "In the domain of intimacy, how many of the 12 items had been endorsed by the female partner? ___ (please fill in)". We postponed this manipulation check to avoid any sensitization of participants to the hypothesized role of base-rate information in the completion of the primary dependent measures. Finally, participants furnished some demographical data. Upon completion of the dependent measures, participants were debriefed, thanked, and dismissed.

Results and Discussion

Manipulation Check. The successful extraction of skewed base rates from a series of observations is a necessary condition for the emergence of PCs. In the present study, participants had to extract the base rates of 'yes' responses for each target person in each of four domains, and a reasonable test for the accuracy with which this premise had been fulfilled consists in the subjective ecological correlation between partners' endorsement base rates. Subjective ecological correlation coefficients were computed from participants' estimates of the base rates of 'yes' responses provided by the female partner and by the male partner, respectively, in the four domains of life. Overall, the pattern of base-rates had been extracted accurately from the stimulus series (see the top row of Table 1). The grand mean of .85 is sufficiently high to warrant the assumption that participants noticed that partners' endorsement base rates jointly varied as a function of domain of life, and 2 x 2 ANOVA of the subjective ecological correlations with the between-participant variables presentation (alternate vs. block-wise), agreement (item-wise vs. scale-wise) revealed no effects, all $F(1, 60) < 1$, *ns*.[2]

Indirect contingency measures. Indirect measures of the subjectively perceived correlation were derived from extrapolation task and cued recall performance by computing the delta p entailed in participants' 'yes' and 'no' responses at the aggregate level. (i.e., the probability of his endorsing items she had endorsed minus the probability of his endorsing items she had rejected). The mean aggregate level delta p produced in the extrapolation task can be taken from the second row of Table 1. As expected, the mean perceived contingency consistently exceeded the actual correlation in all conditions, as is evident from series of *t*-tests against the objective value of .2, conducted separately for each experimental condition (for the block-wise/item-wise condition $t(15) = 6.32$, $p < .001$; for the block-wise/scale-wise condition $t(15) = 3.08$, $p < .01$; for the alternate/item-wise condition $t(15) = 1.50$, $p < .10$; for the alternate/scale-wise condition $t(15) = 4.15$, $p < .001$).[3] A 2 x 2 ANOVA of the extrapolation task data with the between-participant variables presentation and agreement revealed no significant effects (all $F(1, 60) < 1.20$, *ns*. All of the mean values found using a sample of 64 real participants fell into the interval of the simulation's mean +/- 1 standard deviation. Thus, the computer simulation approximates the empirical findings far better than the alternative standards of comparison (i.e., the domain level correlation, the aggregate level correlation, or the ecological correlation).

The mean aggregate level delta p computed from the cued recall data further substantiates the position that the hypothesized base-rate reproduction strategy can account for the magnitude of the perceived correlation in the indirect measures. As can be seen in the third row of Table 1, all means fell into the range of .41 to .55, and a 2 x 2 ANOVA of the cued recall data with the between-participant variables presentation and agreement revealed no significant effects (all $F(1, 60) < 1.67$, *ns*). Moreover, a series of *t*-tests against the objective value of .2, conducted separately for each experimental condition revealed that all mean values deviated from the actual aggregate level contingency in the predicted direction (all $t > 3.11$, $p < .01$).

Recall accuracy. Our predictions rest on the assumption that the complexity of the stimulus material should counteract the coordination of partners' responses to the individual items of the questionnaire. Yet, the block-wise presentation of the items pertaining to each of the four subscales used for half the design might have helped participants detect the actual contingencies in the stimulus series. Therefore, we calculated the correlation between participants' cued recall responses and partners' actual responses as a measure of cued recall accuracy (see the bottom row of Table 1). All means fell into the range between .55 and .60, and a 2 x 2 ANOVA of cued recall accuracy with the between-participant variables presentation and agreement revealed no significant effects, all $F(1, 60) = < 1$, *ns*. Thus, cued recall accuracy was impressive overall, and it did not vary meaningfully with experimental conditions.

Experiment 2

Experiment 2 aimed at an extended replication of our initial demonstration, using a refined procedure. First, we raised the proportion of diverging observations to 50 percent in each domain, in order to set the predicted PC clearly apart from the contingency information entailed in the cell frequencies. As a consequence, both the resulting domain level contingency of -.33 and the aggregate level contingency of 0 are distinct from positive PCs (see the bottom panel of Figure 1). Note in passing that the elevated proportion of diverging responses also yields a less extreme ratio of frequent over rare responses of only 3:1 in each domain of life, thus adding to the generalization of the effects reported in Experiment 1.

Second, we administered twice as many items in each domain of life during the extrapolation task, in order to enhance the reliability of the main indirect measure. To obtain a reliable estimate of the magnitude of perceived correlations based on base-rate reproduction, we also ran another simulation programmed to apply the reproduction strategy to the modified extrapolation task. For 1000 simulated participants, the simulation randomly assigned the frequent answer in a given domain to 6 items and the infrequent answer to 2 items, thus reproducing the base rates of frequent and infrequent answers. If both partners had endorsed most items in a domain, that is, a random generator produced a 'yes' for her in response to six new items in a first run – as well as for him in a second run. And if both partners had rejected most items in a domain, a 'yes' was produced for her in response to two new items in a first run – and for him in a second run. Figure 3 gives the percentages of cases falling into different intervals of the aggregate level delta p are given in. The mean delta p of the distribution is .24, and the standard deviation is .15.

Third, we wanted to examine whether the repeated application of the base-rate strategy may translate into explicit expectancies linking frequent and rare responses to each other. To this end, participants were asked to provide explicit estimates of the subjective conditional probability of his endorsing items she had endorsed and of his endorsing items she had rejected in each domain of life. In addition, participants were asked to indicate the degree to which they used several response strategies while working on the indirect measures. Finally, we explored a manipulation intended to affect the ease with which participants could coordinate partners' endorsing versus denying responses to the individual items in the extrapolation task. Half the participants indicated her responses and his responses to the additional items on the same form, whereas participants indicated her responses and his responses on two separate forms as in Experiment 1.

Method

Participants and design. A total of 40 undergraduate psychology students (31 women and 9 men; *M*(AGE) = 23.25, *SD*(AGE) = 5.82) were recruited for a study on clinical judgment in partial fulfillment of a course requirement. Under an equal *n* constraint, participants were randomly assigned to one of the four experimental conditions resulting from the orthogonal variation of the between-participants variables extrapolation task (simultaneous vs. successive) and relationship quality (good vs. bad). Several contingency measures were administered within participants.

Procedure. The same general procedure was used as in Experiment 1, with the following exceptions. First, all within-domain correlations were clearly negative and the aggregate correlation was set to zero. Second, the number of items per domain in the extrapolation task was doubled. In an attempt to manipulate the ease with which participants could coordinate partners' endorsing versus denying responses to the individual items, moreover, two different versions of the extrapolation task were administered. Participants in the simultaneous condition indicated her responses and his responses on the same form, whereas participants in the successive condition indicated her responses and his responses on two separate forms. Third, in order to examine whether the perceived contingencies expressed in the indirect measures would translate into explicit expectations, estimates of the conditional probabilities of matching vs. non-matching responses were included as a direct measure of the perceived contingency. Specifically, participants provided conditional percentage estimates (a) for the likelihood of his endorsing items she had endorsed and (b) for the likelihood of his endorsing items she had denied. Note that these percentage estimates correspond to the proportion of observations falling in the left cell of the upper row and of the lower row, respectively, in the 2 x 2 table of a given domain. Subtracting the latter estimate from the former thus yields an analog of delta p. Finally, participants also completed several ratings assessing the utilization of different strategies in working on the indirect contingency measures. Using six-point rating scales from 1 (completely incorrect) to 6 (completely correct), participants indicated the degree to which they agreed with four statements, each of which described one of the following strategies: (1) reproduction of the endorsement base rate for each partner and domain, (2) reproduction of the proportion of converging answers for each domain, (3) invariant utilization of the modal value for each partner and domain, and (4) pseudocontingent alignment of frequent and rare answers within each domain. The item wordings are given in the Appendix.

Results and Discussion

Manipulation Check. One participant in the simultaneous/bad condition had to be dropped from data analysis, because he failed to recognize that the two partners' base rates of 'yes' responses were substantially correlated. For the remaining participants, however, the pattern of base-rates had been extracted accurately from the stimulus series again (see the top row of Table 2). A 2 x 2 ANOVA of the subjective ecological correlations with the between-participant variables extrapolation task (simultaneous vs. successive) and relationship quality (good vs. bad) revealed a marginally significant main effect of relationship quality, $F(1, 38) = 3.21$, $p < .10$, but no other effects. An inspection of Table 2 reveals that this effect was due to lower values obtained in the simultaneous/bad condition. However, the mean value of .83 is still sufficiently high to warrant the assumption that participants did notice that partners' endorsement base rates jointly varied as a function of domain of life.

Indirect contingency measures. The mean aggregate level delta p produced in the extrapolation task can be taken from the second row of Table 2. Again, the mean perceived contingency consistently exceeded the actual correlation in all conditions, as is evident from *t*-tests against the objective value of 0, conducted separately for each experimental condition (for the simultaneous/good condition $t(9) = 8.45$, $p < .001$; for the simultaneous/bad condition $t(8) = 3.22$, $p < .05$; for the successive/good condition $t(9) = 5.99$, $p < .001$; for the successive/bad condition $t(9) = 1.95$, $p < .10$). A 2 x 2 ANOVA of the extrapolation task data with the between-participant variables extrapolation task and relationship quality revealed a significant main effect of relationship quality, $F(1, 38) = 5.37$, $p < .05$, but no other effects. An inspection of Table 2 reveals that this effect was due to the fact that participants in the bad relationship condition produced somewhat lower perceived correlations. More important to our present purposes, however, is the observation that all of the mean values found using a sample of 40 real participants fell into the interval of the simulation's mean +/- 1 standard deviation. Thus, the computer simulation approximates the empirical findings far better than the alternative standards of comparison (i.e., the domain level correlation, the aggregate level correlation, or the ecological correlation).

The mean aggregate level delta p produced in the cued recall task further substantiates the position that the hypothesized base-rate reproduction strategy can account for the magnitude of the perceived correlation in the indirect measures. As can be seen in the third row of Table 2, all means fell into the range of .22 to .26, and a 2 x 2 ANOVA of the cued recall data with the between-participant variables extrapolation task and relationship quality revealed no significant effects (all $F(1, 38) < 1$, *ns*). Moreover, a series of *t*-tests against the objective value of 0, conducted separately for each experimental condition revealed that all mean values deviated from the true aggregate contingency in the predicted direction (all *t*s > 3.75, p < .01).

Direct contingency measure. A direct measure of the subjectively perceived correlation was derived from the conditional probability estimates by subtracting the likelihood of his endorsing items she had denied from the likelihood of his endorsing items she had endorsed separately within each domain. The mean average delta p entailed in the conditional probabilities can be taken from the fourth row of Table 2. As expected, all means fell into the vicinity of the expected value of .25 again. A 2 x 2 ANOVA of the extrapolation task data with the between-participant variables extrapolation task and relationship quality revealed no significant effects, all $F(1, 38) < 1$, *ns*. In addition, a series of *t*-tests against the objective of -.33, conducted separately for each experimental condition revealed that all mean values deviated from the actual mean contingency in the predicted direction (all *t*s > 5.53, p < .01).[4]

Strategy endorsement. 2 x 2 x 4 ANOVA of the strategy endorsement ratings with the between-participant variables extrapolation task and relationship quality and the within-participants variable strategy (base rate vs. convergence vs. mode vs. alignment) revealed a significant main effect of strategy, F(3, XX) = X.XX, p < .001, but no other effects. The strategy effect reflected the two strategies invoking the utilization of base-rate information, base rate and alignment, were endorsed to some degree, whereas the alternative strategies, convergence and mode, were not (*M*(BASE RATE) = 3.59, *SD*(BASE RATE) = 1.46; *M*(ALIGNMENT) = 3.21, S*D*(ALIGNMENT) = 1.36; *M*(CONVERGENCE) = 2.62, *SD*(CONVERGENCE) = 1.16; *M*(MODE) = 2.72, *SD*(MODE) = 1.47). Thus, the impact of base-rate information on performance was not only visible in the indirect as well as in the direct contingency measures, but also in participants' explicit self-reports regarding their strategic approach to contingency tasks.

Recall accuracy. Again, we calculated the correlation between the cued recall data and partners' actual responses as a measure of cued recall accuracy. The mean cued recall accuracy can be taken from the bottom row of Table 2. All means fell into the range between .71 and .77. A 2 x 2 ANOVA of cued recall accuracy with the between-participant variables extrapolation task and relationship quality revealed no significant effects, all $F(1, 38) = 1$, *ns*.

Although cued recall accuracy was impressive overall, it also varied to some degree. This opened up the possibility to examine the degree to which the effects obtained for the extrapolation task (indirect measure) and for the conditional probability estimates (direct measure) vary as a function of the accuracy with which participants managed to reproduce the stimulus information. To this end, we computed the correlations between all dependent measures. These correlation coefficients are given Table 3. Apart from the trivial finding that the bias in the cued recall task diminishes as the cued recall accuracy increases, none of the other relationships reached conventional levels of statistical significance. Put differently, the effects reported so far obtained independently of the accuracy with which participants could reproduce the actual aggregate level correlation.

General Discussion

The present research aimed at demonstrating that the mere reproduction of the base rates of variables can result in pseudocontingencies. Using a partnership-counseling scenario, participants produced positively correlated judgments on both direct and indirect contingency measures, although the objective contingency between partners' responses had been negative within each of four domains of life in an initial learning phase. As anticipated by computer simulations reproducing the objective base rate of ‘yes’ responses independently for each target person and domain, the main dependent measures yielded positive contingency coefficients, which consistently fell into the vicinity of the predicted value of .48 (Experiment1) and .25 (Experiment 2), respectively. Neither cued recall accuracy nor an experimental manipulation that aimed at facilitating the coordination of the female partner’s responses with the male partner’s responses in the main indirect measure yielded any systematic effects. Put differently, the biased judgments obtained independently of the ease with which participants could reconstruct and/or express the objective contingency in the dependent measures. These findings point to the possibility that the repeated reproduction of the accurately learned base rates lead to some kind of self-persuasion into the belief in a moderate positive contingency between partners’ responses, as evident from our findings for explicit conditional probability estimates.

In several respects, the present report advances our understanding of pseudocontingencies. First, the present findings extend the scope of empirical demonstrations of PCs to the domain of multi-context situations. Previous research relied on stimulus series providing participants with the base rates of two variables in one or two contexts, whereas the present findings obtained using a stimulus series providing base rates for an array of contexts. Second, the findings provide converging evidence for the existence of a distinct mediating process. To reiterate, past research had shown that the heuristic alignment of the more frequent levels of two variables can account for the emergence of PCs in the case of bivariate frequency distributions (e.g., Bluemke & Fiedler, 2007; Fiedler & Freytag, 2004; Kutzner, Freytag, Vogel, & Fiedler, 2007) as well as multi-variate frequency distributions (Meiser & Hewstone, 2004). In the present report, the mere reproduction of the base rates predicted participants’ performance in all experimental conditions and for all dependent measures. These findings thus do not only provide another piece of evidence for the pervasive impact of base-rate information on contingency perception. They also highlight that the mediating processes underlying PCs vary with the task at hand.

One point that deserves attention in generalizing the present findings is that we did not find evidence for an impact of the ecological correlation in the stimulus series on the perceived contingency within domains of life. Although we can only speculate about the reasons, it seems plausible that the present experimental situation rendered the ecological correlation less informative in several respects. First, the number of contexts was relatively small. Given only four domains of life, participants had to monitor no more than eight base rates, a number within the boundaries of short-term memory. However, adding no more than one or two contexts may push the number of base rates to be stored beyond a critical level, thus promoting the utilization of alternative sources of information, such as the ecological correlation across contexts. Second, the ecological correlation was perfect, whereas the within-domain contingencies were weak. To the extent that participants realized these differences in magnitude, the informational value of the ecological correlation may have been diminished. Disentangling base-rate reproduction and ecological bias will require the identification of the conditions under which judgments are based on the pattern of base rates observed within contexts versus on the pattern of base rates observed across an array of contexts (see Vogel, Fiedler, Freytag, & Kutzner, 2008).

Concluding Remarks

As we have pointed out repeatedly (e.g., Fiedler et al., 2008), PCs should not be dismissed as some kind of laboratory research artifact. Rather, the necessary conditions for the emergence of PCs are established quite frequently outside the lab, too. For instance, the experimental situation created here has much in common with the circumstances under which we evaluate the similarity among people in everyday life: Similarities in the endorsement base rate of different people may often reflect nothing but the general desirability of the specific domains of life (Who likes doing the dishes anyway?), and information about different persons' endorsement of the individual "items" often accumulates across episodes dispersed over extended periods of time (thus counteracting the coordination of the individual pieces of information). We hope to have demonstrated convincingly that pursuing the seemingly neutral base-rate reproduction strategy under such conditions will result in inflated perceptions of the similarity of people in our immediate social environment.

References

Bluemke, M., & Fiedler, K. (2007). *The fate of pseudo-contingencies in affective priming*. Unpublished manuscript, 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., & Meiser, T. (in press). Pseudocontingencies: A long-overlooked cognitive illusion. *Psychological Review*.

Fiedler, K., Freytag, P., & Unkelbach, C. (2007). Pseudocontingencies in a simulated classroom. *Journal of Personality and Social Psychology*, *92*, 665-677.

Hammond, J. L. (1973). Two sources of error in ecological correlations. *American Sociological Review*, *38*, 764-778.

Kutzner, F., Freytag, P., Vogel, T. & Fiedler, K. (2008). Base-rate neglect as a function of base rates in probabilistic contingency learning. Manuscript under revision, Journal for the Experimental Analysis of Behavior.

Meiser, T. (2007). Contingency learning and biased group impressions. In K. Fiedler & P. Juslin (eds.), *Information sampling and adaptive cognition* (pp. 183-209). Cambridge: 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.

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

Vogel, T., Kutzner, F., Fiedler, K., & Freytag, P. (2008). The impact of base rates on the perception of multi-level problem structures. Unpublished manuscript, University of Heidelberg.

Authors' Note

Peter Freytag, Tobias Vogel, Florian Kutzner, and Klaus Fiedler, Department of Psychology, University of Heidelberg, Hauptstrasse 47-51, 69117 Heidelberg, Germany. The research reported in this article was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) to the first and to the last author. The authors are grateful to Oliver Arnold, Julica Rang, Stephanie Tremmel, and Sarah Weisang for their assistance in conducting this study. Correspondence concerning this research may be sent via electronic mail to *pf@psychologie.uni-heidelberg.de*.

Footnotes

1 Note that the reliability of correlation coefficients computed within domains depends on the number of items. In the experiments reported below, the number of items per domain varied between 4 and 12 in the dependent measures. Given this impoverished data basis, we did not feel entitled to “predict” that these noisy indices will reflect a “perceived zero correlation”, although this prediction can be derived from our argument. As a consequence, we refrain from analyzing perceived correlations at the domain level. Nevertheless, it maybe worth mentioning that the gross mean within-domain correlations fell into the predicted range (in Experiment 1: *M* = .XX, *MIN* = .XX, *MAX* = .XX; in Experiment 2: *M* = .07, *MIN* = -.06, *MAX* = .24).

2 The relationship quality variable did not qualify any of the analyses reported below, and was thus dropped from data analysis.

3 All *t*-tests reported in this manuscript were two-tailed.

4 In contrast to the analysis of the aggregate level contingency expressed in both the extrapolation task and the cued recall task, the appropriate test value of the t-tests in the analysis of the average contingency expressed in the conditional probability estimates was -.33, because estimates were conditional on domain of life.

Appendix

Items assessing the utilization of different strategies in working on the indirect contingency measures (the original German wordings appear in italics).

Base rate strategy

"I checked the 'agree' option at a rate corresponding to the proportion of agreeing answers provided by a partner in a given domain."

"*Ich habe so oft 'stimme zu' angekreuzt, wie es dem Anteil zustimmender Antworten eines Partners im jeweiligen Bereich entspricht*."

Convergence strategy

"I checked the agree option for both partners at a rate corresponding to the proportion of converging answers in a given domain."

"*Ich habe für beide Partner so oft dieselbe Antwort angekreuzt, wie es dem Anteil gleicher Antworten beider Partner im jeweiligen Bereich entspricht*."

Modal value strategy

"I always checked the option corresponding to the more prevalent answer provided by a partner in a given domain."

"*Ich habe immer die Antwort angekreuzt, die der vorherrschenden Antwort eines Partners im jeweiligen Bereich entspricht*."

Alignment strategy

"I checked the same option for both partners if they had shown the same response tendency in a domain."

"*Ich habe für beide Partner dieselben Antworten angekreuzt, wenn beide Partner dieselbe Antwort-Tendenz in einem Bereich hatten*."

Table 1. Mean contingency estimates derived from tasks calling for the recall or for the prediction of partners' responses in the four domains of partnership life (Experiment 1).

	Presentation			
	Alternate Agreement		Block-wise Agreement	
Contingency Measure	Item-wise	Scale-wise	Item-wise	Scale-wise
Ecological Correlation (R)	.79 (.50)	.90 (.12)	.84 (.25)	.88 (.21)
Extrapolation Task (P)	.38 (.47)	.52 (.21)	.53 (.21)	.48 (.36)
Cued Recall (R)	.48 (.16)	.55 (.19)	.41 (.25)	.46 (.34)
Cued Recall Accuracy (R)	.60 (.18)	.55 (.10)	.58 (.20)	.56 (.24)

Note. Numbers in parentheses are standard deviations. All indices were computed by aggregating responses over domains of life. R = recall-based contingency measure; P = prediction-based contingency measure.

Table 2. Mean contingency estimates derived from tasks calling for the recall or for the prediction of partners' responses in the four domains of partnership life (Experiment 2).

	Extrapolation Task			
	Simultaneous		Successive	
	Relationship Quality		Relationship Quality	
Contingency Measure	Good	Bad	Good	Bad
Ecological Correlation (R)	.93 (.12)	.83 (.17)	.96 (.06)	.92 (.08)
Extrapolation Task (P)	.39 (.15)	.26 (.24)	.38 (.20)	.17 (.28)
Cued Recall (R)	.26 (.15)	.25 (.16)	.22 (.19)	.23 (.11)
Conditional Probabilities (R)	.30 (.28)	.27 (.28)	.32 (.36)	.31 (.19)
Cued Recall Accuracy (R)	.77 (.11)	.71 (.21)	.75 (.17)	.72 (.17)

Note. Numbers in parentheses are standard deviations. All indices were computed by aggregating responses over domains of life. R = recall-based contingency measure; P = prediction-based contingency measure.

Table 3. Inter-correlation of the main dependent measures (Experiment 2).

Measure	1	2	3	4	5	6
1 Marginal Frequencies	1					
2 Extrapolation Task	-.18	1				
3 Cued Recall	-.23	.20	1			
4 Conditional Probabilities	-.08	.05	.29	1		
5 Cued Recall Accuracy	.11	-.24	-.57***	-.25	1	
6 Explicit Base-rate Strategy	-.07	.04	.21	-.06	-.23	1

Note. *** $p < .001$.

Figure Captions

Figure 1. Frequency distributions underlying stimulus presentation at the domain level and at the aggregate level in Experiment 1 (top panel) and in Experiment 2 (bottom panel).

Figure 2. Percentage of cases yielding specific values of the aggregate level delta p (*M* = .48, *SD* = .18) in a computer simulation of extrapolation task performance reproducing the base rate of ‘yes’ responses independently for each partner and domain (*n* = 1000) in Experiment 1.

Figure 3. Percentage of cases yielding specific values of the aggregate level delta p (*M* = .25, *SD* = .15) in a computer simulation of extrapolation task performance reproducing the base rate of ‘yes’ responses independently for each partner and domain (*n* = 1000) in Experiment 2.

Figure 1

Experiment 1

	Domain 1			Domain 2			Domain 3			Domain 4			Aggregate	
	Male			Male			Male			Male			Male	
Female	Yes	No		Yes	No		Yes	No		Yes	No		Yes	No
Yes	6	2		0	2		6	2		0	2		12	8
No	2	0	+	2	6	+	2	0	+	2	6	=	8	12
	Δ = -.25			Δ = -.25			Δ = -.25			Δ = -.25			Δ = .2	

Experiment 2

	Domain 1			Domain 2			Domain 3			Domain 4			Aggregate	
	Male			Male			Male			Male			Male	
Female	Yes	No		Yes	No		Yes	No		Yes	No		Yes	No
Yes	6	3		0	3		6	3		0	3		12	12
No	3	0	+	3	6	+	3	0	+	3	6	=	12	12
	Δ = -.33			Δ = -.33			Δ = -.33			Δ = -.33			Δ = 0	

Figure 2

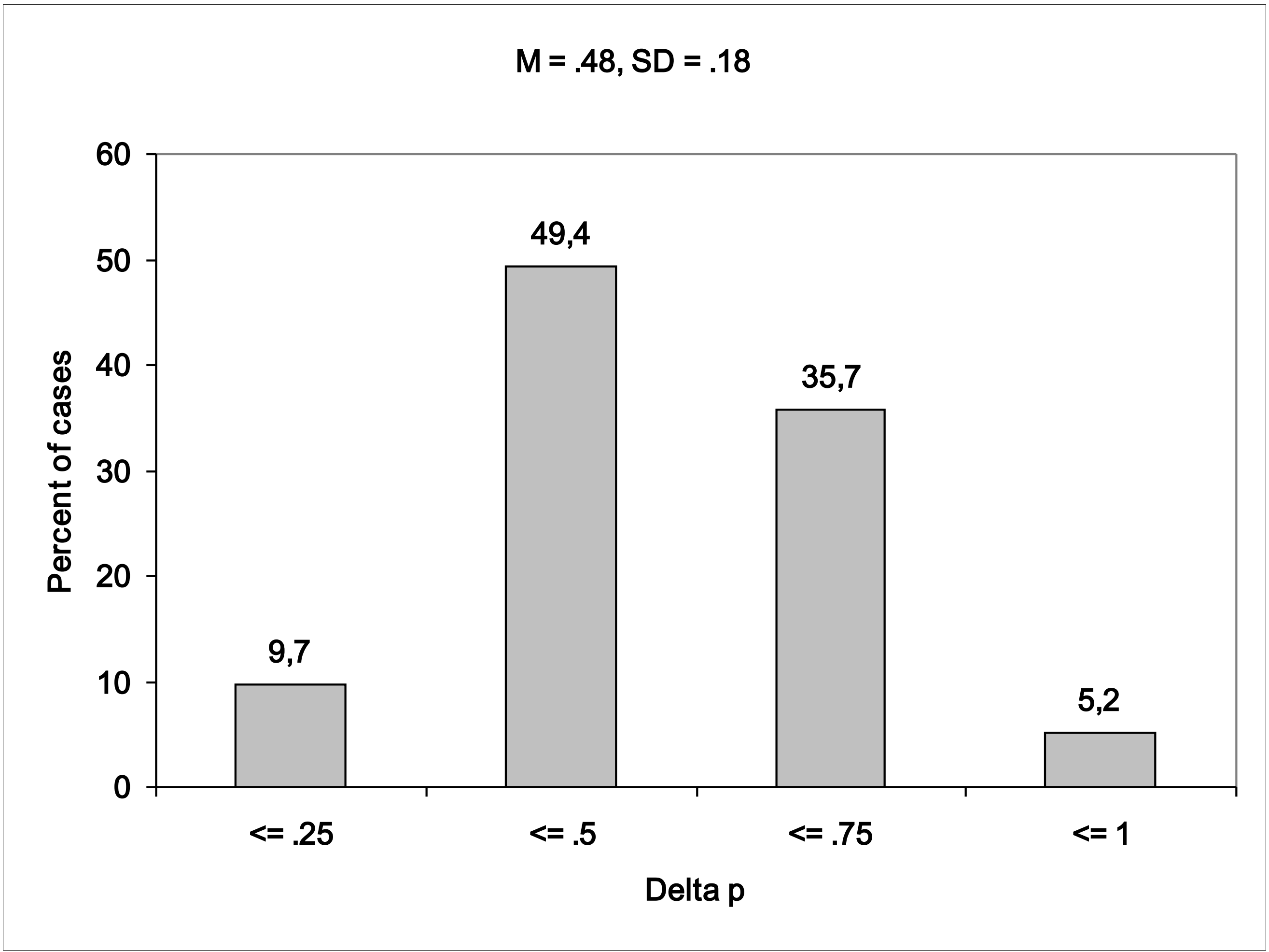
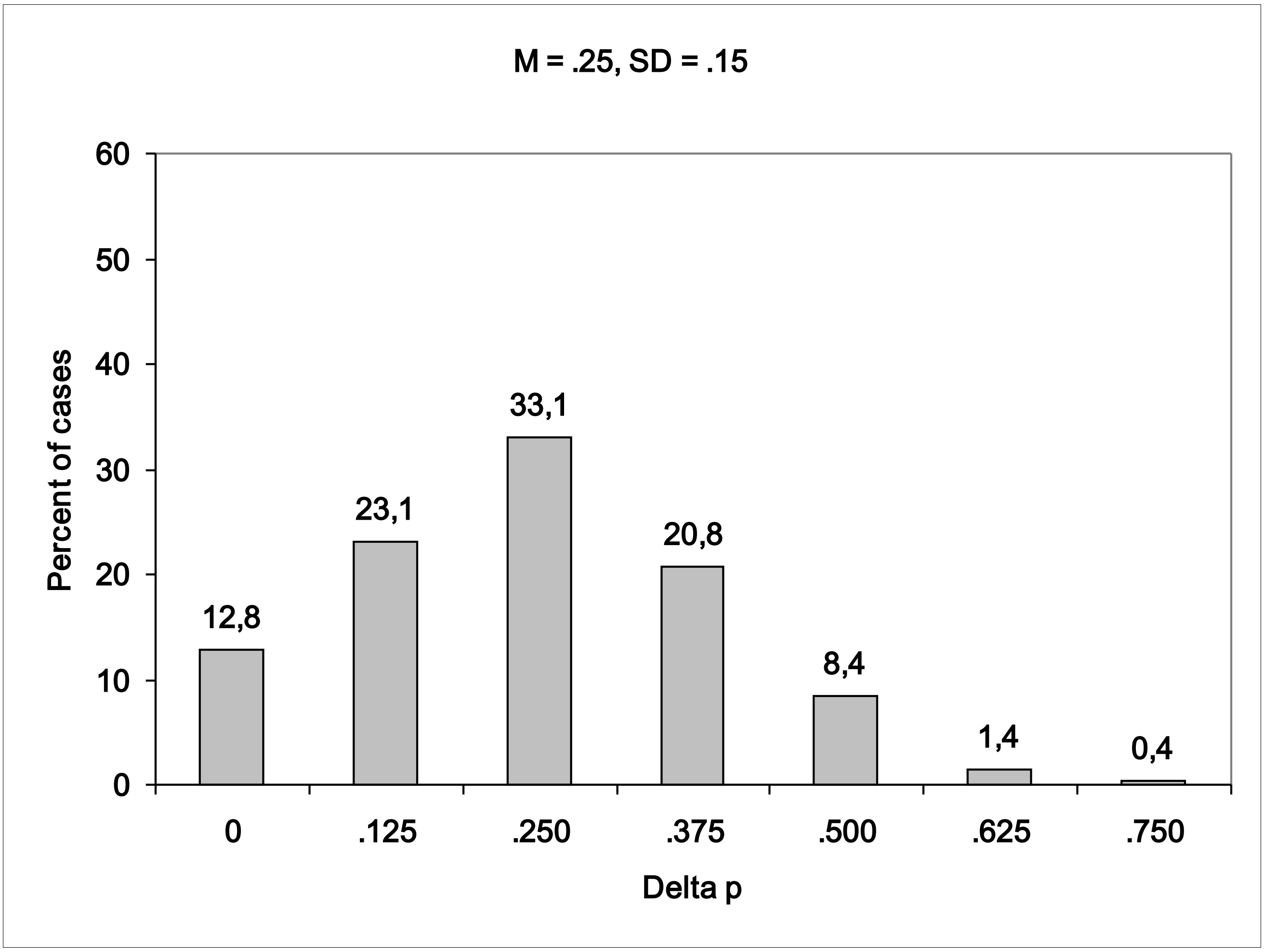

M = .48, SD = .18
60
50
40
30
20
10
0
Percent of cases
49,4
35,7
9,7
5,2
<= .25
<= .5
<= .75
<= 1
Delta p

Figure 3


M = .25, SD = .15
60
50
40
30
20
10
0
Percent of cases
12,8
23,1
33,1
20,8
8,4
1,4
0,4
0
.125
.250
.375
.500
.625
.750
Delta p

Nr.	Author	Title
08-40	Florian Kutzner Tobias Vogel Peter Freytag Klaus Fiedler	Pseudocontingencies in stereotype formation: extending illusory correlations
08-39	Wendelin Schnedler Adam Dominiak	Uncertainty Aversion and Preference for Randomization
08-38	Susanne Abele Sandra I. Vaughan-Parsons Garold Stasser	Information Flow and Influence during Collective Search, Discussion, and Choice
08-37	Peter Freytag Tobias Vogel Florian Kutzner Klaus Fiedler	The reproduction of base-rates promotes pseudocontingencies
08-36	Klaus Fiedler Peter Freytag Thorsten Meiser	Pseudocontingencies: An Integrative Account of an Intriguing Cognitive Illusion
08-35	Alexander Unger Dagmar Stahlberg	The Influence of Self-Regulation on Decision Making
08-34	Dirk Simons Nicole Zein	Kosten aus einer asymmetrischen Informationsverteilung zwischen Abschlussprüfer und Mandant
08-33	Michael Bremert Axel Schulten	The Impact of Supervisory Board Characteristics on Firm Performance
08-32	Jürgen Eichberger Ani Guerdjikova	Case-Based Expected Utility: Preferences over Actions and Data
08-31	Jürgen Eichberger Simon Grant Jean-Philippe Lefort	Neo-additive capacities and updating
08-30	Vladimir I. Danilov Gleb A. Koshevoy Christine Lang	Equilibria with indivisible goods and package-utilities

Nr.	Author	Title
08-29	Thomas Gschwend	Electoral System Change in Belgium 2003: Party Strategies and Voter Responses
08-28	Michael Herrmann	Expectations about Coalitions and Strategic Voting under Proportional Representation
08-27	Karl-Martin Ehrhart Marion Ott Susanne Abele	Auction Fever: Theory and Experimental Evidence
08-26	Michel Regenwetter	Perspectives on Preference Aggregation
08-25	Daniel M. Bernstein Michael E. Rudd Edgar Erdfelder Ryan Godfrey Elizabeth F. Loftus	The Revelation Effect for Autobiographical Memory: A Mixture-Model Analysis
08-24	Markus Glaser Florencio Lopez de Silanes Zacharias Sautner	Looking Inside a Conglomerate: Efficiency of Internal Capital Allocation and Managerial Power Within a Firm
08-23	Markus Glaser Martin Weber	Financial Literacy und Anlegerverhalten
08-22	Markus Glaser Thomas Langer Jens Reynders Martin Weber	Scale Dependence of Overconfidence in Stock Market Volatility Forecasts
08-21	Patrick A. Müller Rainer Greifeneder Dagmar Stahlberg Kees Van den Bos Herbert Bless	The Role of Procedural Fairness in Trust and Trustful Behavior
08-20	Patrick A. Müller Rainer Greifeneder Dagmar Stahlberg Kees Van den Bos Herbert Bless	The Influence of Accessibility Experiences on the Willingness to Cooperate in Negotiations

Nr.	Author	Title
08-19	Jana Janßen Patrick A. Müller Rainer Greifeneder	A field study on the role of ease-of-retrieval in procedural justice judgments
08-18	Rainer Greifeneder Benjamin Scheibehenne	When choosing is difficult: Complexity and choice-overload
08-17	Paul Grout Wendelin Schnedler	Non-Profit Organizations in a Bureaucratic Environment
08-16	Clemens Kroneberg Isolde Heintze Guido Mehlkop	On shoplifting and tax fraud: An action-theoretic analysis of crime
08-15	Hermann Jahnke Dirk Simons	A rationale for the payback criterion
08-14	Peter Dürsch Jörg Oechssler Radovan Vadovic	Sick Pay Provision in Experimental Labor Markets
08-13	Carsten Schmidt Martin Strobel Henning Oskar Volkland	Accuracy, Certainty and Surprise - A Prediction Market on the Outcome of the 2002 FIFA World Cup
08-12	Mathias Sommer	Understanding the trends in income, consumption and
08-11	Hans Gersbach Hans Haller	Club Theory and Household Formation
08-10	Michael F. Meffert Thomas Gschwend	Strategic Voting in Multiparty Systems: A Group Experiment
08-09	Jens Wüstemann Jannis Bischof	Ausweis von Finanzinstrumenten in europäischen Bankbilanzen nach IFRS: Normative Erkenntnisse empirischer Befunde
08-08	Jürgen Eichberger David Kelsey	Are the Treasures of Game Theory Ambiguous?
08-07	Jürgen Eichberger Ani Guerdjikova	Multiple Priors as Similarity Weighted Frequencies