Illusory correlations revisited: The role of pseudocontingencies and working-memory capacity

Andreas B. Eder
Friedrich-Schiller-Universität Jena, Jena, Germany

Klaus Fiedler
Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany

Silke Hamm-Eder
Julius-Maximilian-Universität Würzburg, Heidelberg, Germany

Three experiments investigate illusory correlations in a feedback learning paradigm. Diverging from a standard paradigm, in which stimuli consist of joint observations of group–behaviour pairs, participants were asked to guess the group reference of positive and negative stimulus behaviours. They only knew that one group was larger than the other, and the stimulus series soon revealed that positive (negative) behaviours appeared more frequently in the stimulus series than negative (positive) behaviours. Regardless of whether feedback of the actual group reference was provided or not, the predominant valence was more strongly associated with the large than the small group. This illusory-correlation effect was evident in memory-based measures at the end of the stimulus series as well as in the online predictions during stimulus presentation. The strength of illusory correlations increased with decreasing working-memory capacity, operationalized either by an interpersonal differences measure or a cognitive-load manipulation. The occurrence of illusory correlations in the absence of joint observations about group–valence pairs (in the no-feedback condition and in the early phase of the online prediction task) can be explained as a reflection of pseudocontingency inferences.

Keywords: Illusory correlation; Pseudocontingency; Feedback; Working-memory capacity.

Hamilton and Gifford's (1976) seminal demonstration of an illusory-correlation effect that favours majorities over minorities through purely mnemonic processes, independent of prejudice, sentiment, and group membership, continues to fascinate students and researchers in social cognition and intergroup relations. When the same high proportion of positive behaviours is observed in two groups that only differ in the number of observations presented, then the majority (i.e., the group for which more observations are provided) will be judged more positively than the minority (Fiedler, 1991; Hamilton & Sherman, 1989; Mullen & Johnson, 1990). The relative devaluation of minorities is evident in frequency estimates of positive and negative behaviours observed in both groups, in trait impression ratings, and in the cued recall of the group in

517

which positive and negative behaviours have been observed.

In numerous experiments conducted in different labs, the phenomenon has been shown to be robust and replicable over a broad range of task conditions (Fiedler & Walther, 2004; Mullen & Johnson, 1990). However, while the illusion itself is undisputable, a long-lasting and still ongoing debate revolves around its theoretical explanation. For many years it was assumed that the illusion originates in an enhanced memory for negative behaviours of the minority, supposed to reflect the distinctiveness of this rarest combination (Johnson & Mullen, 1994). However, cogent evidence for a memory advantage of these rare observations has not been found when memory performance was analysed more thoroughly, controlling for alternative influence, response bias, and speed–accuracy trade-off (Fiedler, Russer, & Gramm, 1993; Klauer & Meiser, 2000). Convergent evidence from several studies shows that an illusory bias in favour of the larger of two equally positive groups can arise in the absence of selective memory, as a normal consequence of the fact that there are simply more trials to learn the high positivity rate of the majority than there are trials to learn the same rate for the minority (Fiedler, 1996).

Illusory correlation and working-memory capacity

However, common to both explanations—enhanced memory of the minority's negative behaviour or more complete learning of the majority's positive behaviour—is the implication that the illusion should increase with decreasing memory capacity. If learning and memory are perfect, the prevailing positive tendency should be equally recognized for both groups. Only if memory is imperfect is there latitude for one group to be more strongly associated to the predominant valence. Indeed, Mullen and Johnson's (1990) meta-analysis confirms that the strength of illusory correlations increases with increasing memory load, but decreases under conditions that improve memory. For instance, explicit impression

judgements, statement presentation in tabular rather than sequential form, and individuals rather than groups used as judgement targets were shown to reduce errors in judged covariation, presumably because these conditions enhanced memory and efficient encoding of the presented information in working memory (e.g., Hamilton, Dugan, & Trolier, 1985; Pryor, 1986; Sanbonmatsu, Sherman, & Hamilton, 1987). Conversely, high processing demands like long statement series, non-optimal study times, and explicit valence judgements of the behavioural statements were shown to enhance illusory correlation (Fiedler et al., 1993; Gordon, 1997; Mullen & Johnson, 1990).

However, although clearly suggestive and plausible on theoretical ground, most of this evidence is only indirect and equivocal. Little is known about the precise mechanisms that mediate the influence on memory of such factors as circadian rhythm, list length, or encoding instructions. In some studies the direction of influence was only inferred post hoc from the observed data pattern (e.g., whether an explicit valence judgement of the group behaviour improves or impairs memory performance; Fiedler et al., 1993). Complicating things further, other studies showed that an increased processing load attenuates illusory correlations under certain circumstances, supposed to reduce capacity for deliberative processing (Stroessner, Hamilton, & Mackie, 1992). These mixed results led Spears and Haslam (1997) to conclude that increased as well as decreased processing load may both attenuate illusory correlations.

In the present research, we try to clarify the role of memory capacity in several ways. By including separate measures of working-memory and short-term memory capacity, we demonstrate that only the former construct is theoretically relevant and empirically related to the strength of the illusion correlation bias. To substantiate the causal influence of working memory, we then manipulate working-memory load experimentally, obtaining stronger illusions when a demanding secondary task is introduced. Moreover, we demonstrate a completely new, online variant of illusory correlations. We show that the majority is not only judged more favourably in retrospective

memory-based ratings, but that an illusory group difference is apparent from the beginning, when participants make online guesses of the group associated with every presented stimulus behaviour. Such online guesses do not entail explicit memory of the entire stimulus list, but they nevertheless call for implicit memory of valence−group associations. It is thus possible that working memory moderates both the memory-based and the online measures of illusory correlations (cf. Hastie & Park, 1986). For a theoretical explanation of the process underlying this new online variant of illusory correlations and its dependence on working memory, we relate our findings to a recently discovered cognitive illusion called pseudocontingency (Fiedler, Freytag & Meiser, 2009).

Illusory correlations in a feedback learning task

For an empirical test, we constructed the following feedback-learning task. Each trial of an extended series starts with the presentation of a behaviour description, which is clearly positive or negative. Participants are asked to predict whether the given behaviour comes from Group A or Group B; they know that Group A is larger than Group B. Immediately after their prediction, a feedback informs them whether they were right or wrong, thereby revealing the stimulus behaviour's actual group reference. As in previous experiments, most stimulus behaviours are positive, though we also run a condition in which negative valence prevails. The prevalence of positive behaviours is the same for the majority and the minority group. At the end of the learning series, participants then give frequency estimates and impression ratings for both groups. We expected to find subjective group differences not only in these end-of-sequence judgements (as in previous studies) but also in the online predictions made during the presentation stage.

Supportive evidence for both expectancies would be novel and of theoretical interest. On the one hand, virtually all previous studies have used passive observation tasks, and it is not at all clear whether the illusion will persist when participants engage in continuous, active predictions. It has been argued, indeed, that the phenomenon is eliminated when instructions encourage online inferences as opposed to memorization (Pryor, 1986). To the extent that intergroup behaviour in real life not only is amenable to passive observation but often involves active participation and "foraging", checking the generality of the illusion under such conditions is of interest. On the other hand, demonstrating that group discrimination is already apparent in the participants' online predictions would be a completely novel finding with intriguing theoretical and practical implications.

Why does this online version of illusory correlations afford a theoretical challenge? One answer, as already noted, is that the online prediction task is very low in explicit memory demands. Unlike the memory-based judgements at the end of the stimulus sequence, guessing the group reference of new behaviours requires neither the retrieval of the entire stimulus series nor the retrieval of a specific group−behaviour pair, which has not been presented yet. Predictive guessing merely involves the utilization of an implicit-memory function that is sensitive to the strength of group−valence associations. Although this implicit-memory function demands little cognitive capacity, it may nevertheless be sensitive to working-memory load. As convincingly demonstrated by Kareev and colleagues (Kareev, 1995a, 1995b; Kareev, Lieberman, & Lev, 1997), memory load serves to reduce the "window size" or the effective stimulus sample that organisms use for making statistical inferences. That is, under high cognitive load, inferences may only utilize the last few items or a greatly reduced time sample from the entire list. Because the differences observed in smaller samples (e.g., between Group A and Group B) tend to be larger than those in larger samples, this means that small samples may inform stronger correlation inferences and/or reward-maximizing prediction strategies than large samples, regardless of whether they are correct or incorrect (Fiedler & Kareev, 2006; Gaissmaier, Schooler, & Rieskamp, 2006). Therefore, both explicit estimates as well as online predictions should reflect stronger illusory correlations when working memory is constrained.

But why do we expect an online bias in group discrimination at all, given the strange, if not surrealistic, task to predict the group reference of behaviours under complete uncertainty, well before feedback information provides authentic information about the pairing of groups and behaviours? Note that after 12 stimulus observations, the majority of participants will not have seen a second pairing of negative behaviour with Group B, even though the bias to associate B with more negative valence than A is already apparent, as will be seen shortly. What aspect of the stimulus input, then, should motivate participants to associate Group B with more negative and less positive behaviour? To answer this puzzling question, we have to resort to a new explanatory principle, the notion of pseudocontingencies (PC), which actually implies the phenomenon we are proposing.

Pseudocontingencies

In the context of a binary prediction task, a PC illusion (Fiedler & Freytag, 2004; Fiedler et al., 2009; Fiedler, Freytag, & Unkelbach, 2007) can be understood as an alignment of two base-rate-driven response biases. First, participants know in advance that Group A is more likely to occur than B. Second, it takes only a few trials to recognize that positive valence is more likely than negative valence. Thus, participants quickly learn to expect Group A and positive valence and not to expect Group B and negative valence. In such a situation, a PC illusion leads participants to infer a correlation: The more frequent (expected) Group A seems to co-occur with the more frequent (expected) positive valence, whereas the less frequent (unexpected) Group B seems to co-occur with the less frequent (unexpected) negative valence. This logically unwarranted inference is made even when the actual joint frequencies do not support this conclusion—that is, when the positivity rate of A and B is actually the same, or actually higher for B (cf. Fiedler & Freytag, 2004).

To set this base-rate-driven illusion apart from illusory correlations supposed to reflect the biased processing of joint frequencies, PCs have been shown under conditions that do not provide any genuine correlation information. For instance, when the prevalent value on one variable is learned in one stimulus series, and the prevalent value of another variable is observed in another, separate series, without a chance to coordinate joint observations, a PC will nevertheless be inferred (Fiedler & Freytag, 2004; McGarty, Haslam, Turner, & Oaks, 1993; Meiser & Hewstone, 2006).

PC inferences have been demonstrated in various paradigms using different stimulus materials and dependent measures, including memory-based end-of-sequence judgements, online predictions, and resulting impression ratings (for an overview, see Fiedler et al., 2009). Therefore, a PC illusion can also be expected to produce a bias toward aligning the larger (smaller) group with the more (less) prevalent valence in the present online guessing task, even though feedback about group–valence pairings appears rather late and does not really support for the superiority of A over B. After all, the actual correlation is zero, and many salient, corrective feedback events will even highlight that B is actually not that negative, and corrective feedback for A may highlight that A is less positive than suggested. However, for a crucial test of the PC account, and to rule out any alternative account in terms of a genuine contingency between stimulus valence and groups through sparsely appearing feedback, we also demonstrate an online bias when all feedback information is omitted. We nevertheless expect both online predictions and postsequence judgements to exhibit group discrimination, providing distinct evidence for a PC effect.

Preview of empirical tests and experimental predictions

To investigate the role of memory capacity in illusory correlations, we employed separate measures of short-term-memory (STM) and working-memory (WM) capacity. STM is typically thought to reflect primarily domain-specific storage that involves coding and rehearsal of limited visuospatial and verbal information. WM, instead, is typically construed as a domain-general, multicomponent system that is responsible for active maintenance and organization of

information in the face of ongoing processing and/ or distraction (for overviews see Baddeley, 2003; Engle, 2002). Evidence supporting this distinction stems from studies showing that STM and WM correlate differentially with higher order abilities like reasoning, reading comprehension, and other complex cognitive tasks (e.g., Conway, Kane, & Engle, 2003; Daneman & Carpenter, 1980). Stronger correlations between WM measures and complex cognitive tasks corroborate the view that WM more than STM constrains the processing component involved in inductive inferences (cf. Unsworth & Engle, 2007).

To the extent that the strength of illusory correlations increases with decreasing memory capacity, this should be due to WM rather than STM. On the one hand, restrictions of WM capacity should render the memory-based judgements of frequency and group impressions more inaccurate and noisy, and this impairment should be most pronounced for the minority. Judgements of B should thus be less sensitive to the actual valence differences than judgements of the majority, A. On the other hand, the hypothesized illusion in the online predictions should also increase with decreasing WM capacity, because limited cognitive resources trigger the simple base-rate-driven response biases supposed to underlie PC inferences.

To test these considerations, we conducted three experiments. In Experiment 1, we compared the standard setting of an illusory-correlation experiment with a modified task setting that asked participants to make active online predictions in a feedback-learning task. We expected the group discrimination phenomenon to be evident in both memory-based judgements and online predictions. Moreover, constraints on WM should strengthen illusory correlations. We included distinctive measures for STM and WM, expecting only the latter to moderate the illusion.

In Experiment 2, a secondary task was introduced to manipulate working memory experimentally, rather than relying on tests measuring individual differences. We expected to replicate the double manifestation of illusory correlations in online predictions and memory-based

judgements as well as the impact of working memory on the illusion.

Experiment 3, finally, was an attempt to demonstrate online and memory-based biases when all judgements have to rely on the participants' own predictions of expected group–valence pairings, in the absence of any feedback about the actual associations. In this task setting, which only allows for PC inferences while not providing any genuine correlation information, we nevertheless expected to find systematic online and memory-based biases.

EXPERIMENT 1

In Experiment 1, individual differences in STM and WM capacity were separately assessed with a Digit Span Test and an Operation Span Test, respectively. Both span tasks have proven to be reliable and valid measures of memory functions (Conway et al., 2005). We expected that reduced working-memory capacity, as distinguished from short-term-memory capacity, should lead to stronger illusory correlations.

Method

Participants and design
A total of 48 volunteers (33 women), most of them students, participated in the experiment either for course credit or for a small monetary gratification (1 euro). The sample had a mean age of 26.9 years ($SD = 9.0$). All participants were native speakers of German. One half of them were randomly assigned to a condition with online predictions in the presentation phase; the other half were assigned to a presentation mode condition with standard presentations of the behavioural statements. In addition to this between-subjects factor, target groups and the valence of behavioural descriptions varied within participants.

Materials
Individual differences in memory capacity were assessed with a Digit Span Test and with an Operation Span Test (Turner & Engle, 1989). In

the Digit Span Test, participants were to recall a number of digits in correct serial order. In each trial, a number of digit words were successively flashed for a brief time period (100 ms), centred from left to right in a row, on the screen. After 2 s the participant was cued by a signal to pronounce the digits in the order of appearance. The test started with a presentation of three digit numbers, and the number of presented digits was individually adjusted according to a staircase procedure: when the digits of a list were correctly recalled in serial order, the next list was one item longer; if incorrect, it was one digit shorter (to a minimum of two digits). This procedure was repeated until the registration of 10 shifts from a correct recall to an incorrect one or vice versa. The median of the number of items that were correctly recalled before and subsequent to an error indexed the STM span (Mueller, Seymour, Kieras, & Meyer, 2003; Schweickert, Guentert, & Hersberger, 1990).

WM capacity was assessed with a variant of the Operation Span Test that showed good psychometric properties in a German validation study (Hamm, 2002). In each trial, a string of simple arithmetic operations was presented together with a suggested answer and with a to-be-remembered word on the right side—for example, "$(4 \times 3) - 5 = 6$? BANK". The participants' task was to respond verbally whether the answer following the equal sign was true or false and then to say the word that followed the operation. A time limit of 15 seconds was given for these responses, and then the next trial was initiated. After varying sets of two to six of these operations, three question marks prompted the participant to recall the words of the previous set in the order of their appearance. In total, three sequences containing one set of each size were presented in randomized order. Half of the suggested answers to the arithmetic operations were correct; half were incorrect. Sets that were reproduced in correct order and with all arithmetic operations correctly verified were scored as correctly reproduced. The sizes of correctly reproduced sets were then summed to form an index of WM capacity (Conway et al., 2005).

For the illusory-correlation experiment, 32 statements describing favourable behaviours (M = 6.1, SD = 0.12) and 16 statements describing unfavourable behaviours (M = 1.7, SD = 0.9) were selected from a standardized pool according to their evaluative norms on a 7-point scale (Ehrenberg, Cataldegirmen, & Klauer, 2001). In addition, 48 male first names were selected and were randomly assigned to the behavioural descriptions to form behavioural statements like "David takes the problems of others seriously". A total of 24 positive statements and 12 negative statements were randomly assigned to Group A; the remaining 8 positive and 4 negative statements described members of Group B. Thus, the ratio of positive to negative behaviours within each group was 2:1, constituting a zero correlation between behaviour favourability and group membership. In the prediction condition, the behavioural statements were presented without indication of the group membership (see example above). In the standard presentation condition, the statements contained a reference to the group membership of the described person (e.g., "David from Group A takes the problems of others seriously").

Procedure

Participants completed first the Digit Span Test and then the Operation Span test. They were then informed that a number of behavioural statements would be presented on the monitor, describing members of two different groups. They were told that Group B was markedly smaller than Group A. The described persons were mentioned to be just a random selection of the members of each group. In the standard presentation condition, the participants' task was to memorize and retain the presented information as well as possible (cf. Pryor, 1986). The behavioural statement was shown for 8 s at the centre of the monitor, followed by a blank period for 1,500 ms (cf. Klauer & Meiser, 2000). Then the next statement appeared. In the online prediction condition, the participants' task was to guess the group membership of the described person without time pressure. In each trial, a behavioural statement about a person appeared, and the participant predicted his group membership by pressing the

letter key "A" or "B" of the keyboard. When the prediction was correct, the word "RICHTIG!" (correct) appeared in green for 500 ms at the centre of the screen; when the prediction was incorrect, the word "FALSCH!" (false) was shown in red for the same time period. Thereafter, the correct group membership appeared at the centre of the screen. Participants were informed in advance that they would earn 3 eurocents for each correct prediction. The total amount of earned money was displayed in the right upper corner. The next trial was initiated by pressing the spacebar.

After the presentation of all 48 statements, participants were asked to estimate the frequency of undesirable behavioural descriptions in each group. This frequency estimation was done separately for Group A and Group B. Participants were given the total number of statements about each group (36 and 12) and were asked to estimate how many statements from each group described undesirable behaviours. Frequency estimations were entered with the keyboard without any time pressure. At the end of the session, participants were asked for biographical data, were thanked and debriefed, and were paid for participation.

Results

Memory measures

Averaged across all participants, memory span was 6.6 ($SD = 1.4$, range: 4–9) in the Digit Span Test and 15.1 ($SD = 6.0$, range: 3–26) in the Operation Span Test. The correlation between both measures was low ($r = .25$), corroborating that different aspects of short-term memory performance were measured in both tests. The two presentation conditions did not differ in their memory performance on both tests, with both Fs < 1.

Illusory correlation

Participants' guesses of the group membership in the online prediction task and the frequency estimates in the memory-based judgement task were used to construct 2×2 contingency tables representing the numbers of positive and negative behaviours for each of the two groups. A phi-coefficient indicating the extent of judged covariation was derived for each of these individual contingency tables and was transformed into a Fisher z score (cf. Hamilton & Gifford, 1976).

In the online-prediction condition, the average participant predicted Group A for 24.17 (out of 32) desirable behaviours, which corresponds to 75.5%. In contrast, when undesirable behaviour statements were presented, the rate of Group A predictions decreased to 56% (8.96 out of 16). Consequently, the z-transformed phi-coefficients derived from these online prediction rates were significantly larger than zero ($M_{ZPhi} = 0.23$), $t(23) = 3.34$, $p < .01$.

It should be noted that the bias in favour of Group A was already apparent in the online predictions for the first 12 stimulus trials ($M_{Phi} = 0.28$), $t(23) = 3.08$, $p < .01$.[1] In these trials, 70.7% and 42.0% of all positively and negatively described persons were assigned to Group A, respectively, highly overestimating the negativity of Group B. This early bias suggests that the illusion is independent of the presentation of a reasonable number of joint observations of groups and valence.

In the online-prediction condition, an illusory-correlation effect also emerged in the subsequent memory-based frequency estimates; z-transformed correlations computed from the frequency estimates of negative behaviours ($M_{ZPhi} = 0.19$) were significantly larger than zero, $t(23) = 4.30$, $p < .01$. Mean estimates of the relative proportion of negative behaviours in each group were much higher for Group B (58.3%) than for Group A (38.6%), even though the actual proportion of negative behaviours was identical in both groups (33.3%).

In the standard presentation condition, the proportion of estimated negative behaviours in each

---

[1] In these few trials, the predictions of 1 participant in Experiment 1 and the predictions of 3 participants in Experiment 2 yielded a phi-coefficient of 1.0 that could not be transformed into a Fisher's z score. Therefore, untransformed phi-coefficients were entered in these analyses.

group did not differ for majority ($M_A = 43.9\%$) and minority ($M_B = 45.8\%$), resulting in a non-significant illusory correlation effect ($M_{ZPhi} = 0.02$), $t < 1$. The difference on this measure between the online prediction and standard presentation conditions was significant, $t(46) = 2.82$, $p < .01$. Thus, the biased online inferences during the prediction task might have strengthened the illusion obtained in subsequent frequency judgements.

Memory influences

The influence of memory limitations on illusory correlation was analysed in multiple regression analyses with presentation mode and the scores of the digit-span test and the operation-span test as predictor variables and the z-transformed phi-coefficient computed from the frequency estimates as criterion. This analysis revealed a significant regression weight of presentation mode ($\beta = .36$, $p < .01$), reflecting the illusory correlation effect in the online-prediction condition and a significant regression weight of the operation-span scores ($\beta = -.27$, $p < .05$, one-tailed). The negative sign indicates that participants with lower WM capacity perceived minority group members less positively than participants with a higher WM span. The regression coefficient for the digit-span test scores was negligible ($\beta = -.05$, $t < 1$), showing that STM capacity was unrelated to illusory correlations.

In an analogous regression analysis using phi-coefficients derived from online predictions as a criterion, neither digit-span scores ($\beta = -.18$) nor operation test scores ($\beta = .09$) predicted the strength of illusory correlations significantly (both ts < 1). The corresponding zero-order correlations are $-.165$ and $.057$, respectively.

Discussion

As expected on theoretical grounds, the first experiment provided evidence for the contention that an illusory bias in favour of the larger of two groups that exhibit the same high prevalence of positive behaviours is already apparent in online predictions of the group membership, given the valence

of the behavioural description as a cue and given that the Group A base rate is higher than the Group B base rate. Well before genuine correlation information about joint observations of groups and valence could have been assessed, participants utilized the easily apparent base rates to predict a majority membership more frequently than a minority membership. Participants predicted the less prevalent group particularly often when the less prevalent negative behavioural statement was given. This online manifestation of an illusory bias was actually more robust than the illusory correlation obtained in memory-based frequency estimates. The failure to replicate the normal effect in the standard-presentation condition is unexpected and at variance with prior studies. However, some null findings have always been observed (Mullen & Johnson, 1990), though rarely published, and illusory correlation effects are frequently obtained only in one out of several effect measures. In the absence of an explanation, we tend to attribute this unexpected result to the fact that the inclusion of a prediction task may have sensitized participants to the unequal base rates. As a consequence, an illusion in the traditional memory-based measures was confined to the prediction condition.

EXPERIMENT 2

Whereas WM capacity was assessed as an individual-difference measure in the first study, in the second study we included an experimentally controlled manipulation of memory load. WM capacity was manipulated using the random interval repetition technique (RIR; Vandierendonck, De Vooght, & Van der Goten, 1998) that selectively disrupts the central executive but not the slave systems (phonological loop and visuospatial sketchpad). All participants conducted the online-prediction task during stimulus presentation. Only half of them conducted the RIR as a secondary task supposed to induce cognitive load. At the end of the experimental session, all participants provided evaluative group-impression ratings, a cued-recall test of the group associated with

positive and negative behaviours, and estimates of the frequency of negative behaviours shown by the two target groups. We expected the following pattern to cross-validate our interpretation of the first experiment: (a) Illusory biases favouring Group A over Group B should be obtained in both online measures and memory-based measures; (b) cognitive load should increase the strength of the obtained illusory correlations.

Method

Participants

A total of 66 volunteers (47 women; mean age 25.1 years, $SD = 5.8$) participated in the experiment either for course credit or for a small monetary gratification (1 euro). All participants were native German speakers; none of them had participated in Experiment 1.

Stimuli, design, and procedure

All participants had to perform group membership predictions in the statement presentation phase, and the behavioural statements presented were identical to those of Experiment 1. Half of the participants performed random taps (RIR) during the statement presentations; the other half viewed the behavioural statements without a secondary task. The sequence of trial events in the condition with RIR was as follows: After a blank period of 2,000 ms, a countdown from 3 to 2 to 1 was displayed on the screen, taking a total time of 1,050 ms (350 ms for each number). For the time period thereafter, participants were instructed to respond to the occurrence of a brief beep signal (20 ms) as quickly as possible with a single tap of the left mouse button. The interval between successive beeps was either short (500 ms) or long (800 ms), with random selections between both intervals until a time period of 2,500 ms was covered. Then the behavioural statement was presented in the centre of the screen for further 2,500 ms, with random tapping continued. After this time period, the screen cleared, and the tapping task stopped. Participants were prompted to enter their guess of the group membership without time pressure, and feedback on the

correctness of the prediction and the actual group membership was given just like in Experiment 1. The next trial was initiated with a press of the space bar. In the condition without RIR, the sequence of trial events was identical, but without a secondary tapping task. The countdown was replaced with a blank screen, and the timing of the events was identical to that of the RIR condition.

Subsequent to the presentation of all 48 statements, measures of group perception were collected in the following order: (a) evaluative group ratings, (b) a cued-recall assignment task, and (c) the frequency estimation task. In the evaluative rating task, six questions asked for separate evaluative ratings of Groups A and B on a scale ranging from −4 to +4. One question asked for an overall impression of the Group A (B) descriptions (very positive−very negative), one for a sympathy judgement towards Group A (B) members (very dislikeable−very likeable), and one for an appraisal of how pleasant a contact with Group A (B) members would be (very unpleasant−very pleasant). Each question was shown in the centre of the computer screen with a 9-point scale displayed below, and the evaluative rating was entered with a mouse button press on the respective scale value box without time pressure.

In the cued-recall assignment task, the 48 behavioural statements were presented again in randomized order, and participants were to recall whether the statement described a Group A or a Group B member. To prevent a memory cueing by the first name, the name was replaced with "person XY" in the behavioural statement (e.g., "Person XY takes the problems of others seriously"). The group assignment was entered with the buttons "A" and "B" without time pressure. The frequency estimation task was the same as that in Experiment 1 and was presented last to preclude carry-over effects of self-generated frequency estimates on subsequent measures.

Results

Participants' guesses of the group membership in the online prediction task, their frequency

estimates of undesired group behaviours, and their trait assignments were used to construct 2×2 contingency tables representing the number of positive and negative behaviours for each of the two groups. A phi-coefficient was derived for each of these individual contingency tables and was transformed into a Fisher z score. The scores of the three evaluative ratings of Group A (Cronbach's $\alpha = .91$) and Group B (Cronbach's $\alpha = .87$) were recoded to scale values from 1 to 9 and then averaged to form a single index of the group evaluation. Table 1 shows the mean guessing frequencies in the online prediction task, the frequency of behaviour assignments in the cued-recall task, and the frequency estimates of group behaviours as a function of behaviour type (positive, negative) and group origin.

Tapping performance
On average, 413.5 bleeps were presented (range: 404–423). Multiple taps (i.e., two or more taps

to a single bleep) were performed in 14.8% ($SD = 10.8$) of the tapping trials. Mean reaction time of correct taps was 285 ms ($SD = 45.8$).

Illusory bias in online predictions
Participants in the load condition, who had performed the RIR task, exhibited a strong illusory correlation effect ($M_{ZPhi} = 0.37$), $t(32) = 4.99$, $p < .01$. In the no-load condition without RIR, the illusory-correlation effect in the online measure was slightly reduced ($M_{ZPhi} = 0.22$), $t(32) = 3.09$, $p < .01$, but the difference between load conditions was not significant, $t(64) = -1.51$, $p > .10$.

Again, the differential bias was already visible during the first 12 stimulus trials, well before the joint occurrence of groups and valence could have been reasonably assessed. A strong group bias was observed in both conditions, with RIR ($M_{Phi} = 0.28$), $t(32) = 3.48$, $p < .01$, and without RIR ($M_{Phi} = 0.30$), $t(32) = 3.44$, $p <$

Table 1. *Influence of working-memory load on illusory correlation in Experiment 2*

		Memory load			
		With (n = 33)		Without (n = 33)	
Online group prediction (12 trials)	Pos A (%)	66.8		74.4	
	Neg A (%)	36.6		45.1	
	Phi (z)		0.28**		0.30**
Online group prediction (all trials)	Pos A (%)	75.8		75.7	
	Neg A (%)	45.1		58.7	
	Phi (z)		0.37**		0.22**
Trait assignment	Pos A (%)	75.3		72.8	
	Neg A (%)	53.8		64.8	
	Phi (z)		0.29**		0.10
Frequency estimation	Neg A (%)	33.4		41.0	
	Neg B (%)	57.3		51.0	
	Phi (z)		0.24**		0.10*
Evaluative rating	Mean A	6.04		5.61	
	Mean B	4.45		5.03	
	Diff		1.59**		0.58

Note: Ratios of online predictions and trait assignments are relative to the total number of positive and negative statements (32 and 16, respectively); ratios of frequency estimations are relative to the number of members in Group A and Group B (36 and 12, respectively).
*$p < .05$. **$p < .01$. Level at which mean is different from zero.

.01, with no reliable difference between both conditions ($t < 1$) (see footnote 1).

Illusory correlations in memory-based judgements
The basic illusion and the moderating impact of WM restrictions were regularly visible in the three offline judgements obtained at the end of the session. For the frequency estimations, a t-test for independent samples revealed stronger illusory correlation effects in the load condition ($M_{ZPhi} = 0.24$) than in the no-load condition ($M_{ZPhi} = 0.10$), $t(64) = 2.21$, $p < .05$. In both groups, though, the z-transformed correlation score was significantly larger than zero, $t(32) = 5.09$, $p < .01$, and $t(32) = 2.32$, $p < .05$, respectively.

In an analogous analysis of the phi-coefficients derived from the cued-recall assignment task, the influence of WM load was also significant in a one-tailed test, $t(64) = 1.77$, $p < .05$. Participants under WM load showed a pronounced illusory correlation effect ($M_{ZPhi} = 0.29$), $t(32) = 3.09$, $p < .01$, whereas the condition without RIR did not ($M_{ZPhi} = 0.10$), $t(32) = 1.61$, $p > .10$.

A comparison of the mean evaluative ratings of Groups A and B revealed more positive evaluations of the majority group than of the minority group (see Table 1). The rating difference was more pronounced with WM load than without WM load, $t(64) = 1.80$, $p < .05$ (one-tailed). Comparisons against zero corroborated that the mean difference in the evaluative group ratings was reliable in the WM load condition ($M = 1.59$), $t(32) = 3.84$, $p < .01$, but not in the no-load condition ($M = 0.58$), $t(32) = 1.52$, $p > .10$.

All three memory-based measures together (i.e., z-transformed evaluative rating differences and z-transformed phi coefficients derived from frequency estimates and cued recall) clearly discriminated between the load and no-load conditions, according to a mixed multivariate analysis of variance (MANOVA), $F(1, 64) = 4.47$, $p < .05$. Separate Hotelling T^2 tests corroborated that illusory correlations were significantly larger than zero in the load condition, $T^2 = 25.98$, $F(3, 30) = 8.12$, $p < .001$, but not in the no-load condition, $T^2 = 5.68$, $F(3, 30) = 1.78$, $p > .10$.

Finally, a single index of the memory-based bias was formed by averaging the z-transformed scores of all three memory-based measures. This memory index was entered with the z-transformed phi-coefficients derived from the online predictions as a variable into a mixed ANOVA, with WM load condition serving as a between-subjects factor. The analysis yielded a main effect of WM load, $F(1, 64) = 4.27$, $p < .05$, and a significant main effect of illusory-correlation measure due to a stronger bias in the online measure than in the combined memory-based measures, $F(1, 64) = 11.28$, $p < .01$. The lack of an interaction between both factors ($F < 1$) suggests that WM load exerted a similar influence on both measures.

Discussion

Experiment 2 corroborated the basic findings using an experimental manipulation of WM load rather than test scores of individual difference. Again, illusory correlations were observed both in the memory-based end-of-sequence ratings and in the online predictions. The strength of the illusion increased with memory load, both in the end-of-sequence judgements and in the online predictions. In any case, it should be kept in mind that WM capacity, as measured by the Operation Span Test and manipulated by the RIR, must not be confused with substantial stimulus memory. What amplified illusory correlation is not the selective recall of specific behaviour−group associations but the impairment of central cognitive resources required to integrate prior information and new stimuli in judgement formation.

Our demonstration of an online bias in the early prediction phase, well before substantial stimulus information had been gathered, was interpreted as a PC effect—that is, as an alignment of two response biases, to expect the large group and the more frequent valence, and to expect the small group and the less frequent valence. The cognitive process underlying PC inferences only relies on unequal base rates but not on joint occurrences of groups and behaviours. However, although the bias found for the first 12 trials is consistent with this notion, an influence of joint observations

cannot be ruled out because more and more group−behaviour pairs can be reconstructed from the accruing feedback information. To completely rule out this source of information as a possible cause of the bias, and to cogently demonstrate a PC-type inference process, we finally ran another experiment in which participants predicted the group reference of stimulus behaviours but never received any feedback.

EXPERIMENT 3

How could such an experimental task be devised? We invented a cover story saying that our research was concerned with impression formation from minimal information. Based on very few behaviours about two groups, participants tried to predict the group association of a longer series of stimulus behaviours. Because they were told that one group appeared more frequently, and because one valence appeared more frequently in the list, the minimal conditions for a PC inference were met. An illusory bias to associate the larger group more than the smaller group with the predominant valence should be apparent both in the online predictions as well as in the final end-of-sequence ratings.

Method

Participants

A total of 42 students (27 women; mean age 23.8 years, $SD = 4.3$) volunteered for a monetary gratification up to maximum of 1.5 euros (see Procedure). One half of the participants were randomly assigned to a condition with twice as many favourable behaviour descriptions as unfavourable ones; the other half were assigned to a condition in which undesired behaviours outnumbered desired ones with the same ratio (i.e., 32:16). All participants were native speakers of German, who were naïve to the purpose of the study.

Materials

In the condition with desirable statements outnumbering undesirable ones, the same stimuli

were used as those in Experiment 1. For the condition with a preponderance of undesirable behaviours, we used the 16 most extreme positive behaviours plus an additional 16 descriptions of undesirable behaviours selected from the standardized stimulus pool by Ehrenberg et al. (2001). Statements and presentation lists were constructed in the same way as in the online prediction condition of Experiment 1, except that the there was a 2:1 ratio of undesirable to desirable behaviours in each group.

To induce initial group impressions, separate sets of 8 (4) desirable and 4 (8) undesirable behavioural statements were randomly combined with a male first name taken from a separate set of 12 names to form behavioural statements about a group member (e.g., "Peter from Group A did not take care of his sick mother"). When favourable behaviours were more frequent than unfavourable behaviours, 4 positive and 2 negative statements referred to both Group A and Group B. In the other condition, each group was described by 2 positive and 4 negative randomly allocated behaviours.

Procedure

The experiment was advertised as a study of judgements that are based on minimal social information. Participants were informed that they would first see a short series of behavioural statements that described randomly selected members of two different groups (Groups A and B). They should carefully read each statement to get a first impression about both groups; no mention was made that one group was smaller or larger than the other one. The previewing of group descriptions consisted of 12 behaviour statements; half of them referred to members of Group A and the other half to members of Group B (see Materials above). The statements were presented in random order, and participants had unlimited time to read each description. The presentation of the next statement was initiated with a press of the spacebar.

After the previewing phase, participants were informed that a longer series of statements would be presented next but this time without indication

of the group membership. Their task would be to predict the group membership of the described person on the basis of the given information. Information was now given that Group B is markedly smaller than Group A. Participants were advised to exploit their initial impression for the group guesses and that 3 eurocents would be earned for each correct prediction.

The procedure of the online prediction task was identical to that of Experiment 1 with following exceptions: First, each statement was presented for 4 s, followed by a screen that prompted a press of the "A" or "B" button of the keyboard. After the response, the next statement was then presented after a blank period of 300 ms. Second, no feedback was given on the correctness of the group guess or the actual group membership of the described person. Participants did not learn the total amount of earned money until the end of the session.

When they had predicted the group association of all 48 stimulus behaviours, participants provided frequency estimates of how many undesirable behaviour descriptions they ascribed to each group (for procedural details, see Experiment 1). Finally, they were asked for biographical data, were thanked and debriefed, and were paid the total amount of money they had earned.

Results and discussion

Participants' guesses of the group membership in the online prediction task and their frequency estimates of undesirable group behaviours were used to construct 2×2 contingency tables representing the number of positive and negative behaviours for each of the two groups. For each of these individual contingency tables a phi-coefficient was derived, the positive (negative) sign of which indicates a more (less) favourable impression of the majority (Group A) and a less (more) favourable impression of the minority (Group B). Table 2 shows the mean guessing frequencies in the online prediction task and the frequency estimates of group behaviours as a function of behaviour type (positive, negative) and group origin, separately for

Table 2. *Illusory-correlation effects in Experiment 3*

| Preponderance of statements | N | Online group prediction (all trials) | | | Frequency estimation | | |
		Pos A (%)	Neg A (%)	Phi	Neg A (%)	Neg B (%)	Phi (z)
Positive	21	65.8	39.9	0.26*	43	65.1	0.24*
Negative	21	47.9	59.8	−0.12	65.1	48.8	−0.17*

Note: Ratios of online predictions and trait assignments are relative to the total number of positive and negative statements; ratios of frequency estimations are relative to the number of members in each group.

*$p < .05$. **$p < .01$. Level at which mean is different from zero.

conditions with predominantly positive and negative behaviours.

Illusory bias in online predictions

Group guesses of 2 participants resulted in a phi-coefficient of 1.0 that could not be transformed into a Fisher's z score. Therefore, untransformed phi-coefficients were used in the following analyses. A t test with independent samples revealed that different illusory-correlation effects were produced in the conditions in which positive and negative behaviours dominated the statement series, $t(40) = 2.18$, $p < .05$. When desired behaviours were more frequent than undesired behaviours, a standard illusory-correlation effect was observed ($M_{Phi} = 0.26$), $t(21) = 2.39$, $p < .05$. When undesired behaviours were more frequent, the mean of the phi-coefficients was negative ($M_{Phi} = −0.12$), but not significantly different from zero, $t(21) = −0.87$, $p > .10$. The unreliable illusory correlation effect in the condition with more prevalent negative statements matches the more general finding of Mullen and Johnson's (1990) meta-analysis that the basic illusory correlation effect is reduced with a prevalence of negative statements.

Illusory bias in frequency estimations

Individual phi-coefficients were z-transformed and then subjected to analyses. Like in the online predictions, different illusory-correlation effects were observed in the conditions in which positive or negative behaviour statements were more frequent,

$t(40) = 3.10$, $p < .01$. With a preponderance of positive behaviours, a standard illusory-correlation effect emerged ($M_{ZPhi} = 0.24$), $t(21) = 2.24$, $p < .05$, indicating a less favourable impression of the minority than of the majority group. In contrast, when the persons were described more frequently with negative behaviours than with positive ones, the effect was reversed in sign ($M_{ZPhi} = -0.17$), $t(21) = -2.21$, $p < .05$, indicating a more favourable impression of the minority than of the majority group.

GENERAL DISCUSSION

The present research provided direct support for the contention that working-memory capacity, as distinguished from short-term-memory capacity, moderates the impact of frequency-based illusory correlations. The tendency to provide higher estimates of the rate of negative behaviours for the smaller of two otherwise equivalent groups increased with decreasing scores on the Operation Span Test. Moreover, an experimental manipulation of cognitive load that selectively disrupted WM functions (cf. Vandierendonck et al., 1998) increased the illusory correlation between valence and groups in all three memory-based measures, providing direct evidence for a causal influence of processing-related memory functions on the emergence of illusory correlations.

The notion that the illusory devaluation of minorities as compared to majorities is mediated by memory functions is shared by proponents of different theoretical accounts of the illusion. Regardless of whether researchers have proposed a substantive recall advantage for negative minority behaviours supposed to be most distinctive or prominent (Hamilton & Sherman, 1989), or more complete learning of the association between groups and the high positivity rate for the majority, due to more learning trials (Fiedler, 1996; Meiser & Hewstone, 2006), the resulting difference can explain illusory correlations in frequency estimates, impression judgements, and cued recall.

However, the role of memory in the generation of illusory correlations was almost always understood in terms of selective learning and memory of substantial stimulus contents. It was assumed that the association of the predominant behavioural valence with the two groups is learned more completely for the majority, or that the association of the minority with the more infrequent behaviour has a recall advantage. While these assumptions about differential learning and memory can account for the basic phenomenon, they cannot account for the moderating impact of WM capacity. The present experiments are the first to demonstrate that WM moderates the strength of the illusion, quite independent of how well the association of groups and valence can be assessed and memorized. Apparently, when the operational functions of WM are impaired, participants have to resort to simple response biases that facilitate illusory correlations.

The basic illusion as well as the moderating influence of WM impairment was not only apparent in the usual memory-based measures at the end of the experimental session (i.e., frequency estimates, cued recall, and impression ratings). Rather, a similar pattern was already observed during the stimulus presentation phase, when participants were asked to guess the group origin of stimulus behaviours. Well before they could gather a reasonable number of observations for each group−valence combination, their predictive guesses associated the larger group with the predominant valence and, by complement, the smaller group with the less frequent valence. This was even the case when no feedback was provided about the actual pairing of behaviours and groups, so that all inferences had to rely on the alignment of base-rate-driven response biases (i.e., to expect the large group and the frequent valence and not to expect the small group and the infrequent valence).

Such an inference strategy—predicting the most prevalent criterion value more frequently for the most prevalent predictor value—could well reflect a PC illusion (Fiedler & Freytag, 2004; Fiedler et al., 2009). The alignment of high base rates of positive behaviours and of Group A (and of

negative behaviours and Group B) leads to the inference that the two high-base-rate events and the two low-base-rate events are correlated even in the absence of joint observations.

Inferences based on PCs are more frugal and less demanding than genuine correlation inferences. It is not necessary to receive and encode joint occurrences of two variables (valence and group); instead, it is sufficient to recognize which level on either variable is more prevalent (in this case, positive behaviours and Group A). This cognitive inference process is fundamentally different from correlation assessment. It only relies on the base rates of separate variables rather than joint frequencies. Technically speaking, whereas correlations use the stimulus frequencies within the cells of a contingency table, PC inferences only use the marginal distributions.

Both findings conveyed in the present research report, concerning the role of working memory in memory-based illusions, and concerning the new version of a relatively independent process leading to a similar illusion, have obvious theoretical and practical implications. The conditions leading to illusory correlations appear to be much less restrictive, and more general, than expected. Unequal base rates alone appear to be sufficient for biases against minorities, or for an advantage granted for the self, one's in-group, or any other category for which information is available at high density. Moreover, impairment of working-memory functions, conceived either as a stable personality disposition or as a consequence of task demands or distractor influences, can increase the strength of the illusory-correlation effect.

REFERENCES

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769–786.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*, 547–552.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, *19*, 450–466.

Ehrenberg, K., Cataldegirmen, H., & Klauer, K. C. (2001). Valenz und Geschlechtstypikalität von 330 Verhaltensbeschreibungen: Eine Normierung für studentische Stichproben [Valence and gender typicality of 330 behavior descriptions: A standardization for student samples]. *Zeitschrift für Sozialpsychologie*, *32*, 13–28.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23.

Fiedler, K. (1991). The tricky nature of skewed frequency tables: An information loss account of distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology*, *60*, 24–36.

Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environments. *Psychological Review*, *103*, 193–214.

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

Fiedler, K., Freytag, P., & Meiser, T. (2009). *Pseudocontingencies: An integrative account of an intriguing cognitive illusion*, Manuscript submitted for publication.

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

Fiedler, K., & Kareev, Y. (2006). Does decision quality (always) increase with the size of information samples? Some vicissitudes in applying the law of large numbers. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *32*, 883–903.

Fiedler, K., Russer, S., & Gramm, K. (1993). Illusory correlations and memory performance. *Journal of Experimental Social Psychology*, *29*, 111–136.

Fiedler, K., & Walther, E. (2004). *Stereotyping as inductive hypothesis testing*. New York: Psychology Press.

Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 966–982.

Gordon, R. A. (1997). The moderation of distinctiveness-based illusory correlation: The impact of circadian variations and personal need for structure. *The Journal of Social Psychology*, *137*, 514–526.

Hamilton, D. L., Dugan, P. M., & Trolier, T. K. (1985). The formation of stereotypic beliefs: Further evidence for distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology*, *48*, 5–17.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*, 392–407.

Hamilton, D. L., & Sherman, S. J. (1989). Illusory correlations: Implications for stereotype theory and research. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski, & W. Stroebe (Eds.), *Stereotype and prejudice: Changing conceptions* (pp. 59–82). New York: Springer.

Hamm, S. (2002). AG-Spanne: Validierung eines deutschsprachigen Verfahrens zur Erfassung der Arbeitsgedächtnisspanne. [AG-Spanne: Validation of a technique for assessing the working memory span in the German language]. *Berichte aus dem Psychologischen Institut der Universität Bonn*, *28*(2), 1–23.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, *93*, 258–268.

Johnson, C., & Mullen, B. (1994). Evidence for the accessibility of paired distinctiveness in distinctiveness-based illusory correlation in stereotyping. *Personality and Social Psychology Bulletin*, *20*, 65–70.

Kareev, Y. (1995a). Positive bias in the perception of covariation. *Psychological Review*, *102*, 490–502.

Kareev, Y. (1995b). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition*, *56*, 263–269.

Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General*, *126*, 278–287.

Klauer, K. C., & Meiser, T. (2000). A source-monitoring analysis of illusory correlations. *Personality and Social Psychology Bulletin*, *26*, 1074–1093.

McGarty, C., Haslam, S. A., Turner, J. C., & Oakes, P. J. (1993). Illusory correlation as accentuation of actual intercategory difference: Evidence for the effect with minimal stimulus information. *European Journal of Social Psychology*, *23*, 391–410.

Meiser, T., & Hewstone, M. (2006). Illusory and spurious correlations: Distinct phenomena or joint outcomes of exemplar-based category learning? *European Journal of Social Psychology*, *36*, 315–336.

Mueller, S. T., Seymour, T. L., Kieras, D. E., & Meyer, D. E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1353–1380.

Mullen, B., & Johnson, C. (1990). Distinctiveness-based illusory correlations and stereotyping: A meta-analytic integration. *British Journal of Social Psychology*, *29*, 11–28.

Pryor, J. B. (1986). The influence of different encoding sets upon the formation of illusory correlations and group impressions. *Personality and Social Psychology Bulletin*, *12*, 216–226.

Sanbonmatsu, D. M., Sherman, S. J., & Hamilton, D. L. (1987). Illusory correlation in the perception of individuals and groups. *Social Cognition*, *5*, 1–25.

Schweickert, R., Guentert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science*, *1*, 74–77.

Spears, R., & Haslam, S. A. (1997). Stereotyping and the burden of cognitive load. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 171–207). Oxford, UK: Blackwell.

Stroessner, S. J., Hamilton, D. L., & Mackie, D. M. (1992). Affect and stereotyping: The effect of induced mood on distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology*, *62*, 564–576.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*, 1038–1066.

Vandierendonck, A., De Vooght, G., & Van der Goten, K. (1998). Interfering with the central executive by means of a random interval repetition task. *Quarterly Journal of Experimental Psychology*, *51A*, 197–218.