Journal of Experimental Social Psychology 45 (2009) 1117–1122

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Reports

Mix me a list: Context moderates the truth effect and the mere-exposure effect

Alice Dechêne [a,*], Christoph Stahl [b], Jochim Hansen [a], Michaela Wänke [a]

[a] Department of Psychology, University of Basel, Missionstrasse 60/62, CH-4055 Basel, Switzerland
[b] Institute for Psychology, Albert-Ludwigs-Universität Freiburg, Engelbergerstrasse 41, D-79106 Freiburg, Germany

ARTICLE INFO

Article history:
Received 29 April 2009
Available online 28 June 2009

Keywords:
Truth effect
Mere-exposure effect
List context
Fluency
Repetition

ABSTRACT

When participants are repeatedly presented with an unfamiliar stimulus, this stimulus is rated as more likable (*mere-exposure effect*) or more valid (*truth effect*) as compared with a similar non-repeated stimulus. Both effects have been discussed as effects of fluency. Typical research designs on these effects involve a test phase in which ratings of both repeated and non-repeated stimuli are required. Based on research on moderators of fluency effects, we propose that the procedure of assessing the effects with mixed lists of repeated and non-repeated stimuli contributes strongly to the emergence of both effects. Two experiments found that the truth effect and the mere-exposure effect were strongly moderated by whether mixed lists or only repeated items were used at the test phase: whereas strong effects occurred in a context of repeated and non-repeated stimuli, the effects vanished with only repeated stimuli. Methodological and theoretical implications are discussed.

The *mere-exposure effect* and the *truth effect* are constants in introductory social psychology and consumer behavior textbooks (e.g., Aronson, Wilson, & Akert, 2002; Felser, 2001). Repeatedly presenting previously unfamiliar stimuli increases liking for that stimuli (*mere-exposure effect*), and repeatedly presenting unfamiliar statements increases the likelihood that they are believed (*truth effect*). Both of these repetition effects are well-replicated and have been found under various conditions. There is abundant evidence that repeated exposure makes stimuli more likeable (e.g., Bornstein, 1989; Zajonc, 1968) and subjectively more credible (e.g., Bacon, 1979; Begg, Anas, & Farinacci, 1992; Brown & Nix, 1996; Hasher Goldstein, & Toppino, 1977).

However, a look at the literature suggests that both effects may be closely tied to the experimental procedure. In nearly all studies, participants were presented with mixed lists of old and new stimuli in the test phase (heterogeneous lists). A recent meta-analysis on the truth effect (Dechêne, Stahl, Hansen, & Wänke, submitted for publication) found only a single study that used a homogeneous list of only old statements at test (Schwartz, 1982); in this study, the truth effect emerged only for one set of items but not for another, and the effect size was rather small. Thus, it is possible that the type of test list may influence the truth effect. Likewise, the typical procedure for mere-exposure studies uses test lists of old and new stimuli. Indeed, one study that used a homogeneous test list failed to find a mere-exposure effect (Zajonc, Swap, Harrison, & Roberts, 1971). In sum, the evidence for a truth effect and a mere-exposure effect in homogeneous lists is rather weak. We suggest that the lack of evidence for both effects in homogenous lists is no coincidence but reflects a systematic contribution of the procedure—the use of mixed lists—to the emergence of the effect. Both effects may be enhanced by heterogeneous test lists, and diminished or even eliminated when homogeneous test lists are used. The present studies investigated the effect of heterogeneous vs. homogenous test lists on both effects.

There is indeed a theoretical argument why the composition of the test list (heterogeneous vs. homogeneous) may moderate the truth and mere-exposure effects. Both effects are explained as results of processing fluency, which is defined as the experienced ease of ongoing cognitive processes (e.g., Whittlesea, 1993). Repetition enhances the ease of subsequent processing (Johnston, Dark, & Jacoby, 1985). This cognitive feeling is (mis)attributed to liking (Bornstein & D'Agostino, 1994; Reber, Winkielman, & Schwarz, 1998) or subjective truth (Reber & Schwarz, 1999; Unkelbach, 2007). Evidence for the fluency account comes from studies in which ease of processing was manipulated directly. For example, Reber and Schwarz (1999) presented statements either with high or low color contrasts, and found that high-contrast (i.e., fluent) statements were rated more probably true than low-contrast (i.e., non-fluent) statements. Likewise, high-contrast (i.e., fluent) drawings received higher prettiness ratings than low-contrast (i.e., non-fluent) drawings (Reber et al., 1998). More recent research found the use of fluency in judgments to be moderated by its discrepancy from a standard (Hansen, Dêchene, & Wänke, 2008; Hansen & Wänke, 2008; Westerman, 2008; Whittlesea & Leboe, 2003; Whittlesea & Williams 1998, 2001a, 2001b; Willems & Van der Linden, 2006). This discrepancy-account holds that a deviation from an expected or experienced standard drives fluency

* Corresponding author. Fax: +41 (0) 61 2670628.
E-mail address: alice.dechene@unibas.ch (A. Dechêne).

0022-1031/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2009.06.019

effects. By definition, a signal needs to deviate from the surrounding context; likewise, the experienced fluency needs to feel different from other experiences in order to be meaningful. Whittlesea and Williams (1998) demonstrated that perceived discrepancy of an actual fluency experience against an expected standard is responsible for higher feelings of familiarity. Similarly, the ease with which relevant arguments can be generated affected subsequent attitudes only when the experience deviates from a standard (Hansen & Wänke, 2008). Furthermore, fluency-based illusions of recognition memory were more powerful when fluent stimuli were encountered in a context of less fluent stimuli (Westerman, 2008). Applied to truth effects, Hansen and colleagues (2008) extended Reber and Schwarz (1999) and showed that the truth effect for high- compared to low-contrast statements was more pronounced when the current stimulus' contrast level was discrepant from previous stimuli. Likewise, Willems and Van der Linden (2006) found that the mere-exposure effect was moderated by the expectations regarding the fluency of the stimuli. Taken together, these findings suggest that fluency effects are stronger when a stimulus' processing fluency feels discrepant from the fluency experienced by processing other stimuli in the same context.

This is precisely the condition provided by mixed or heterogeneous lists but not by homogenous lists. In a mixed list, repeated and new stimuli are presented in the test phase. This implies that one subset of test stimuli—the repeated ones—is more fluent than another subset—the new stimuli. Thus, the mixed list represents a situation in which the fluency experienced while processing repeated stimuli feels discrepant from the fluency experienced while processing new stimuli. In a homogenous list, where all stimuli are repeated, all stimuli feel (more or less) equally fluent. As a consequence, test lists consisting of only repeated stimuli should result in less pronounced truth- and mere-exposure effects than mixed lists. The potential impact of mixed vs. homogenous lists was investigated for truth effects in Experiment 1 and for mere-exposure effects in Experiment 2.

Experiment 1

In Experiment 1, ambiguous trivia statements were rated for their subjective truth in two sessions. We varied, in session 2, whether the repeated statements were presented together in a list with new statements (*mixed-list*), or whether solely old (*all-old*) or new (*all-new*) statements were presented.

The truth effect has been computed in two ways: (a) as the increase in truth ratings for repeated statements between sessions 1 and 2 (*within-items criterion*); and (b), as the difference between truth ratings for repeated vs. new statements in session 2 (*between-items criterion*). We predicted the increase in subjective truth between sessions 1 and 2 (i.e., the within-items effect) to be more pronounced in a mixed list than in an all-old list. We also expected the advantage for old items compared to new items (i.e., the between-items effect) to be more pronounced in a mixed list as compared to when both were presented separately to different groups (i.e., the all-old vs. the all-new conditions).

Two experiments were conducted with different sets of stimuli at two different universities. Except for these varieties in stimuli and location, both are identical in design and procedure; they are reported together.

Method

Participants and design

Eighty-four undergraduate psychology students (approx. 61% female) from two European universities participated voluntarily on two successive occasions, separated by a one-week interval. Different sets of statements (sets 1 vs. 2) were used at the two universities.

We implemented a three (List-Type at session 2: all-old vs. mixed-list vs. all-new) by two (Material: set 1 vs. set 2) between-subjects design. Participants were randomly assigned to one of three list conditions. In the *mixed-list condition* participants were presented with a mixed list of repeated and new statements in session 2. The *mixed-list condition* reflects the commonly used design (e.g., Hasher, Goldstein, & Toppino, 1977), and accordingly we expect a truth effect on both the within-items and the between-items criterion, replicating classical effects,

In the *all-old condition*, at session 2, participants rated the same statements as in session 1. We expect a smaller within-items effect in this condition because the statements' higher fluency should not be experienced as discrepant against a standard, and should thus be less likely to be used as a cue for the truth judgment.

In the *all-new condition*, participants rated new statements in session 2. This condition served as a control condition. We expect no difference in truth ratings between sessions because, in both sessions, participants rated non-repeated (i.e., non-fluent) statements.

Materials and procedure

We used 36 trivia statements from various domains. Half were actually true (e.g., "Marion Morrison gained fame under his stage name John Wayne"), half were actually false (e.g., "The University of Toulouse is the world's oldest university and was built in 1121"). The subjective truth of all statements had been rated between 3.5 and 4.5 on a 7-point scale in a pretest. Therefore, statements are relatively ambiguous with regard to their subjective truth. True and false statements did not differ initially ($M_{true} = 4.02$, $M_{false} = 3.98$, $F < 1$). Statements were randomly divided into four sets (A, B, C, and D) of nine statements each, with mean pretest ratings between $M = 3.96$ and $M = 4.02$ ($F < 1$; $SD = 1.20$–1.37), and comparable length ($M = 9.8$–11.1 words). Two sets were randomly selected as the critical sets and two served as control sets (see Table 1).

Participants were tested together in the classroom. The statements were presented in a questionnaire. Participants rated the truth of each statement on a 7-point scale (1 = *definitely false,* 7 = *definitely true*). One week later the procedure was repeated with the above-mentioned three conditions. Each session lasted about 10 min.

Results

Within-items criterion

To test our hypothesis that a truth effect is moderated by the nature of the test list, we compared the within-items effect (i.e., the difference in truth ratings of items presented in both sessions) obtained in the all-old vs. the mixed-list condition. We also included the Material factor. That is, we conducted a two (Session: 1 vs. 2) by two (List Context: all-old vs. mixed-list) by two (Material: set 1 vs. set 2) ANOVA on the truth ratings for critical repeated items, with repeated measures on the first factor. The results show a main effect for session, $F(1, 52) = 10.26$, $p < .01$: items were rated as more probably true in session 2. As expected, this effect was qualified by a Session by List Context interaction, $F(1, 52) = 9.29$, $p < .01$. Follow-up analyzes revealed that, in the mixed-list condition, repeated statements were rated more probably true in session 2 ($M = 4.30$, $SD = .65$) than in session 1 ($M = 3.75$, $SD = .56$), $t(27) = 3.65$, $p = .001$. Thus, as expected, a large truth effect ($d = .88$) emerged in a mixed context including old and new statements. This was not the case in the all-old condition. Here, subjective truth did not increase from session 1 ($M = 3.84$, $SD = .60$) to session 2 ($M = 3.85$, $SD = .45$), $t < 1$, $d = .02$. The results are shown in Fig. 1.

Table 1
Subsets of items used as critical vs. control statements in sets 1 and 2.

	All items repeated condition				Old and new statements mixed condition				All items new condition			
	Session 1		Session 2		Session 1		Session 2		Session 1		Session 2	
	Critical	Control	Critical	Control	Critical	Control	Critical	Control	Critical	Control	Critical	Control
Set 1	A	C	A	C	A	C	A	B	A	C	B	D
Set 2	B	D	B	D	B	D	B	A	B	D	C	A

Although not predicted, the increase in subjective truth was so strong in the mixed-list condition that, for set 1, repeated items in session 2 were even rated more probably true than repeated items in the all-old condition, $F(1, 24) = 7.33$, $p = .01$. This pattern was replicated for set 2, although not significantly so, $F(1, 30) = 1.99$, $p = .17$. Because of this difference, the material factor revealed an interaction with List Context, $F(1, 52) = 4.30$, $p < .05$. Importantly, the Material factor did not interact with the predicted effects (i.e., the main effect for session, and the interaction between session and list context), both $Fs < 1$.

Between-items criterion

We hypothesized that repeated items should be rated more probably true than new items in the mixed-list condition. A two (Item Type: old vs. new) by two (Material: set 1 vs. set 2) ANOVA with repeated measures on the first factor was conducted for truth ratings in the mixed-list condition. As predicted, repeated items ($M = 4.30$, $SD = .65$) were rated more probably true than new ones ($M = 3.57$, $SD = .64$) in the mixed-list condition, $F(1, 27) = 15.52$, $p < .001$, $d = 1.13$. In contrast, when presented in two separate lists, truth ratings did not differ between old and new items, $F < 1$. In the all-new condition, there was also no difference between sessions 1 ($M = 3.77$, $SD = .51$) and 2 ($M = 3.81$, $SD = .46$), $F < 1$. No effects of Material were obtained, $F < 1$.

Discussion

We replicated a truth effect when repeated items were presented interspersed with new items in a mixed list. When presented together with new statements, repeated statements' subjective truth increased (within-items criterion); repeated statements were also rated more probably true than new statements (between-items criterion). In contrast, without new statements in the list, neither did an increase in subjective truth occur in repeated statements, nor were repeated statements rated more probably true than new statements that were presented to a different group of participants.

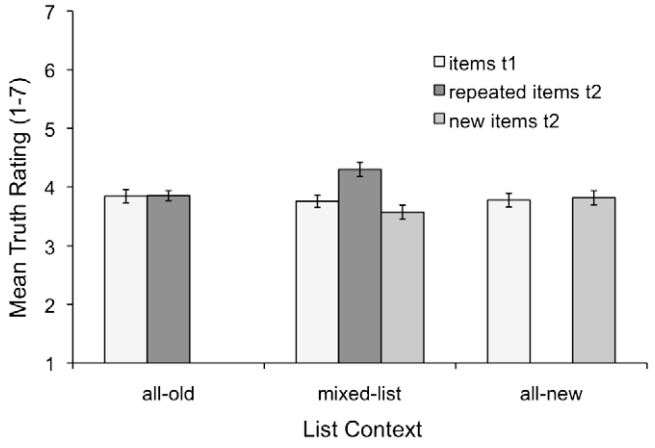

Fig. 1. Mean truth ratings as a function of list context in session 2 (all-old vs. mixed-list vs. all-new) and item status (session 1 vs. repeated items in session 2 vs. new items in session 2). Higher values indicate higher truth ratings. Error bars represent standard errors.

To the best of our knowledge, only one investigation of a truth effect in an all-old context has been published so far (Schwartz, 1982, Experiment 2), and it has been cited in the literature as evidence for the independence of the truth effect from mixed vs. non-mixed presentations of old and new statements in session 2. A closer look, however, reveals the evidence is not all that supportive. In this study, truth ratings increased by 0.07 points on a 7-point scale for one set of items and by 0.03 points on another set of items. Although the effect was significant for the first set (but not for the second set) the effect sizes were rather small, with $d = .14$ for the first set and $d = .06$ for the second set of items.[1] These effect sizes are comparable to our results in the all-old context ($d = .02$), and they reflect, at best, a very small truth effect in the absence of a discrepancy between fluent vs. non-fluent statements (i.e., the homogeneous list context). This was acknowledged by Schwartz, who pointed out that her "[...] finding must be interpreted with caution, however, in view of the weakness of the truth effect found [...]" (Schwartz, 1982, p. 405). In contrast, in the mixed-list context of the present Experiment 1, the effect size for the within-items comparison was $d = .88$, a large effect according to Cohen (1988).

In sum, the data suggest that (increased fluency due to) repetition alone may not suffice to produce the truth effect but that the experience provided by mixed lists contributes significantly to the truth effect. We propose that it is the experience of different levels of fluency, which mixed lists afford, that boosts the truth effect. We will however, discuss alternative accounts in the general discussion session.

Experiment 2

Similar to the truth effect, the mere-exposure effect is discussed as an effect of processing fluency (Bornstein & D'Agostino, 1994). In fact, the mere-exposure effect is also influenced by a discrepancy of expected vs. actual fluency in a perceptual-fluency manipulation (Willems & Van der Linden, 2006). Hence, it seems possible that the fluency-discrepancy inherent in mixed lists may also contribute to the emergence of mere-exposure effects, much as it does for the truth effect. To test this prediction, we manipulated between participants whether repeated stimuli were rated intermixed with new stimuli (*mixed-list*), or whether solely old or solely new stimuli were rated (*separate-list*). Similar to Experiment 1, we hypothesize that the mere-exposure effect is of greater magnitude in the mixed-list as compared to the separate-list condition.

Method

Participants

Seventy-six psychology undergraduates participated in exchange for partial course credit. They were randomly assigned

[1] Effect sizes from that paper were computed from the reported means (Schwartz, 1982, p. 404); the standard deviation was not reported, so we used an estimate (SD=.50) from a recent meta-analysis (Dechêne et al., Submitted for publication).

to the two conditions (mixed-list vs. separate-list); within the separate-list condition, participants were randomly assigned to either the all-old or all-new groups. Ten participants were excluded from analyzes because they reported to be familiar with Kanji symbols (eight participants) or correctly guessed the study's hypotheses (two participants); analyzes including these participants yielded the same pattern of results. The resulting 66 participants (57 female) had mean age $M = 21.81$ ($SD = 3.41$).

Material and procedure

A set of 120 Kanji characters was used. Of these, 40 were randomly selected for presentation for each participant anew. In the mixed-list condition, half of the presented items were randomly selected for the test phase, along with 20 new items. In the separate-list condition, participants were presented with either all of the 40 old items (all-old condition) or with 40 new characters (all-new condition).

The experiment was run in groups of up to four. Each participant was seated in front of a computer. In a first part, participants simply watched the presentation of the stimuli. They saw 40 randomly selected stimuli; each was repeated 10 times in random order. Presentation duration was 13 ms, followed by a 26 ms mask. The inter-stimulus-interval was 4000 ms. In the second part, participants rated their liking (1 = *don't like* to 9 = *like very much*) for each of a randomly ordered set of Kanji characters (20 old and 20 new stimuli in the mixed-list condition; 40 old stimuli in the all-old condition; 40 new stimuli in the all-new condition). Participants were thanked, paid, and debriefed.

Results

We analyzed mean liking ratings for each participant, as well as mean liking ratings for each stimulus. The second set of analyzes allowed us to use repeated-measures analyzes for comparing liking ratings both within as well as between list conditions.[2]

Participants' mean liking ratings were analyzed first. In the mixed-list condition, repeated stimuli were rated more likable ($M = 5.05$, $SD = 1.17$) than new stimuli ($M = 4.74$, $SD = 1.15$); this was significantly so, as revealed by a paired-samples t-test, $t(30) = 2.83$, $p < .05$. In the separate-list condition, ratings were $M = 4.68$ ($SD = 1.02$) for repeated stimuli (all-old condition), and $M = 5.06$ ($SD = 1.24$) for the new stimuli (all-new condition). An independent-samples t-test indicated that this difference was not significant, $t(33) < 1$.

Mean stimulus-based liking ratings were $M = 5.03$ ($SD = 1.25$) for repeated items and $M = 4.77$ ($SD = 1.11$) for new items in the mixed-list condition. For the separate lists, means were $M = 4.64$ ($SD = .84$) for repeated items (all-old condition) and $M = 5.06$ ($SD = .95$) for new items (all-new condition).

A two (List-Type: mixed vs. separate) by two (Exposure: repeated vs. new) repeated-measures ANOVA of stimulus-based mean liking ratings (aggregated across participants) yielded a significant List Type by Exposure interaction, $F(1, 119) = 23.20$, $p < .001$ (Fig. 2). Other effects were not significant. Follow-up analyzes revealed a mere-exposure effect only in the mixed condition: Repeated items were rated as more likable than new items, $t(119) = 2.11$, $p < .05$, $d = .22$. In the separate-list condition, a mere-exposure effect was not found. Unexpectedly, new items were rated as *more* likable than repeated items, $t(119) = 4.04$, $p < .001$.

Comparisons across conditions further revealed that repeated stimuli were rated more likable in the mixed-list than in the separate-list condition, $t(119) = 3.28$, $p = .001$, indicating the importance of discrepant fluency for the mere-exposure effect. Further support for the role of discrepancy comes from the fact that repeated items' liking ratings in the separate-list condition did not differ from new items' liking ratings in the mixed-list condition, $t(119) = 1.04$, $p = .30$. Finally, new items were rated more likable in the separate-list condition than in the mixed-list condition, $t(119) = 2.70$, $p < .01$.

Discussion

The results indicate the predicted mere-exposure effect in the mixed-list condition but not in the separate-list condition. Importantly, the repeated items were rated as more likable in the mixed-list condition than in the separate-list condition. Previous exposure only increased liking when the repeated items were presented interspersed with new items but not when only repeated items were shown. These results are in line with the assumption that an experienced discrepancy in fluency between repeated and new items boosts the mere-exposure effect.

Unexpectedly, new items in the separate-list condition were rated as more likable than the repeated items in that condition, and they were also rated as more likable than new items in the mixed-list condition. Although further research is certainly needed to explain this effect, it is not inconsistent with our assumptions. Much as repeated items may have benefitted from discrepant fluency it is not unlikely that discrepantly "unfluent" items—the new ones—suffer. In other words, mixed lists may emphasize the gap between new and old stimuli, enhancing liking for the old ones and exacerbating liking for the new ones.

Other explanations are also possible: Perhaps participants noticed similarities between some stimuli on the all-new test list and previously shown stimuli. If so, fluency may result from stimulus similarity instead of repetition and may lead to a higher experienced fluency of some similar stimuli (Cleary, Morris, & Langley, 2007). While this explanation would account for the unexpected boost of all-new items it does not provide an alternative account for our main result that repeated items are rated more favorably in the mixed than in the homogenous list.

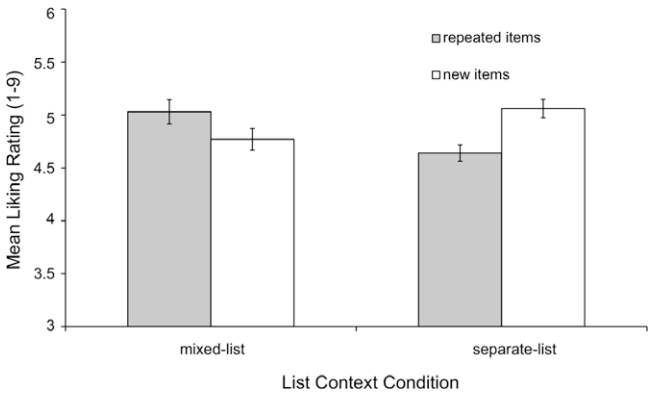

Fig. 2. Items' mean liking ratings as a function of list context (mixed vs. separate) and item type (repeated vs. new). Higher values indicate higher liking ratings. Error bars represent standard errors.

[2] In mere-exposure studies, subject-level repeated-measures analyses are not possible for the separate-list condition because only a single rating is obtained. In mere-exposure studies, stimuli cannot be rated twice—at the second rating occasion, participants would not have been *merely exposed* (but would have rated and thought about stimuli). Therefore pre-experimental baselines cannot be obtained. However, stimulus-level repeated-measures analyses could be conducted because assignment of stimuli to conditions was randomized for each participant anew; thus, each stimulus was presented at least once in each of the conditions. In the following, we performed stimulus-level repeated-measures analyses—instead of between-groups analyses—because of their greater test power.

It may also be argued that the reversed effects in the separate-list condition occurred because participants in the all-old group noticed that all test items were old, and as a consequence, the fluency experience was discounted (Bornstein & D'Agostino, 1992). This account appears problematic for at least two reasons. First, it is unlikely that participants had perfect recognition memory, implying that there were always some items that they considered new. Second, because all participants were presented with the same number of items in the presentation phase and had the same presentation-test delay, recognition memory should not differ between conditions. Therefore, participants should also notice in the mixed-list condition that all old test items were in fact old, and should therefore have discounted their fluency experience in that condition as well. The data clearly shows that this was not the case.

Finally, novelty tends to be perceived as pleasant and interesting (Berlyne, 1970). The items in the all-new condition were all novel, and this novelty might have lead to a more positive evaluation. It remains to be seen whether the unexpected relative increase in likability in the all-new condition can be replicated, and what turns out to be the nature of this effect. Importantly, however, this does not affect the main finding, that repeated items were more likable in the heterogeneous than in the homogeneous conditions.

General discussion

Stirred by a remarkable predominance of using mixed lists for tests of the truth effect and the mere-exposure effect, we conducted two studies to systematically compare effects in mixed and homogeneous lists. In two experiments, we found corresponding results for the truth effect and the mere-exposure effect. Both effects occurred especially under conditions in which repeated and non-repeated stimuli were presented together in a mixed list. Both effects failed to occur when homogeneous lists of only repeated stimuli were presented at time of rating. The little evidence we found in the literature corroborates our results that homogenous lists produce rather weak effects at best (Schwartz et al., 1982; Zajonc et al., 1971). Before we turn to a theoretical explanation of the effect, we would like to point out the methodological implications of this finding.

Repetition effects can be tested in four ways: (a) by comparing responses towards the same stimuli between the first and the repeated exposure within the same perceiver (within-item, within-participant), (b) by comparing responses towards repeated and new stimuli within the same perceiver (between-item, within-participant), (c) by comparing responses towards the same stimuli between pre-exposed participants (for whom the stimuli are repeated) and unexposed participants (for whom the stimuli would be new; within-item, between-participant), and (d) by comparing responses towards repeated stimuli of pre-exposed participants with responses of unexposed participants toward new stimuli (between-item, between-participant).[3] A within-item, within-participant procedure comes closest to the common understanding of mere-exposure or truth effects, namely that stimuli become more attractive or valid *by repetition*. Yet, this proceeding is hardly ever used. For mere-exposure effects, a baseline evaluation contradicts the notion of *mere-exposure* (Bornstein & D'Agostino, 1992). Such a baseline measure would be possible for the truth effect, but as a meta-analysis revealed (Dechêne et al., Submitted for publication), most studies on the truth effect did not obtain baseline ratings. There may be good reasons to rely on between-item-within-participant approaches; for example, a within-participant analysis yields higher statistical power. An item-based analysis of a within-item-between-participant approach (i.e., showing the same items to pre-exposed and to unexposed participants) could be equally powerful (depending on number of items and participants), but as our results suggest, such a presentation of homogenous test lists is less likely to yield repetition effects. Given that mere-exposure and truth effects are such basic phenomena in social psychology it seems noteworthy that both effects are at least to some extent dependent on the method used.

Are truth and mere-exposure effects mere artifacts? Definitely not. Numerous studies including the ones presented here have found that repetition enhances liking and perceived truth, yet the effects may be more confined to specific conditions than has been assumed. We suggest that the emergence of both effects is facilitated when the repeated stimuli are encountered in the context of new stimuli, and the effects are weakened in the context of old stimuli.

This calls for an explanation why the context affects the judgment. We would like to propose an explanation based on the fluency account for both effects, and for the finding that fluency effects are moderated by the difference in processing fluency between the target stimulus and other contextual stimuli. We argued that mixed lists provide a context in which the processing of the repeated stimuli is experienced as discrepantly fluent compared to the new stimuli in the test list. Homogenous lists vary less in the experienced fluency; therefore, the experienced fluency is less likely to be used. It is the deviation from a baseline that makes the fluency signal salient, and thereby a diagnostic cue. This converges nicely with previous research demonstrating that feelings of familiarity result from a discrepancy between actual and expected fluency (Whittlesea & Williams, 1998; Whittlesea & Williams, 2000; Whittlesea & Williams, 2001a; Whittlesea & Williams, 2001b), that illusions of recognition memory due to fluency are more powerful in contexts of less fluent stimuli (Westerman, 2008), that ease-of-retrieval influences attitudes especially when there is a discrepancy between the actual retrieval experience and a previous standard (Hansen & Wänke, 2008), and that truth and mere-exposure effects are more pronounced when the fluency is in contrast to expectations (Hansen, Dechêne, & Wänke, 2008; Willems & Van der Linden, 2006). The present results are in line with this account.

Of course, this does not rule out other explanations. But alternative explanations have to account for the absence of an effect in the all-old list in *both* studies. Given completely different materials, procedures, and dependent variables, the effect is not likely to be material- or procedure-based. As discussed above, we cannot entirely rule out that participants in the mere-exposure study recognized the old items as old in the all-old list and for this reason did not rely on their experience. Crucially, however, this would not explain the results for the truth effect. In fact, if a homogenously old list caused the recognition of some or all of the stimuli as old, truth effects should be more—not less—pronounced. Contrary to mere-exposure effects, which are diminished when the stimuli are recognized as old, truth effects are enhanced when the statements are recognized as old (e.g. Bacon, 1979). Even when participants recognized items as having been presented in the same experiment, truth ratings increased (Bacon, 1979). To us, a discrepancy-fluency account appears to be a parsimonious explanation for both effects. However, future research may provide alternative or additional accounts.

It should be noted that the discrepancy-fluency account does not imply that mere-exposure or truth effects are impossible in homogenous lists. Let us assume that the processing facilitation due to repetition varies among repeated stimuli; not all repeated stimuli are experienced as equally fluent. With increasing variance in fluency, mere-exposure or truth effects are more likely to emerge even when the judgment occurs in an all-old context. In

[3] The latter procedure calls for pretests of the equivalence of the stimuli or counterbalancing, which would make it rather a within-item-between-participant procedure.

particular, in lists where some items have been repeated more often than others, mere-exposure effects can be explained by discrepant fluency, despite the fact that all items have been shown before. Mere-exposure or truth effects may also occur when the context at judgment does not consist of equally fluent stimuli – a situation which most closely resembles natural environments outside of the laboratory.

To conclude, the present findings suggest that a heterogeneous list is an important factor not only for the truth effect but also for the mere-exposure effect. At the least, the present findings demonstrate that list composition can modulate both effects. These findings can help guiding researchers' choice of method and help avoiding erroneous conclusions about the nature and magnitude of the truth effect and the mere-exposure effect.

Acknowledgements

This research was supported by the DIP Project 4.2 "Metacognition: A Window to the Conscious and Unconscious Determinants of Behavior". An earlier version of this article was awarded the "Best Paper Award 2008" by the European Social Cognition Network (ESCON). We thank Klaus Fiedler for providing the opportunity for the data collection. We also thank Florian Kutzner and Mareike de Boer for their help with the data collection.

References

Aronson, E., Wilson, D. W., & Akert, R. M. (2002). *Social psychology* (4th ed.). New Jersey: Prentice Hall.

Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 241–252.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121*, 446–458.

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception and Psychophysics, 8*, 279–286.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin, 106*, 265–289.

Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere-exposure effect. *Journal of Personality and Social Psychology, 63*, 545–552.

Bornstein, R. F., & D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional model of the mere-exposure effect. *Social Cognition, 12*, 103–128.

Brown, A. S., & Nix, L. A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1088–1100.

Cleary, A. M., Morris, A. L., & Langley, M. M. (2007). Recognition memory for novel stimuli: The structural regularity hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 379–393.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.).

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (submitted for publication). The truth about the truth: A meta-analytic review of the truth effect.

Felser, G. (2001). *Werbe- und Konsumentenpsychologie (2 Aufl Advertising and Consumer Psychology)* (2nd ed.). Berlin: Springer.

Hansen, J., Dechêne, A., & Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology, 44*, 687–691.

Hansen, J., & Wänke, M. (2008). It's the difference that counts: Expectancy/experience discrepancy moderates the use of ease-of-retrieval in attitude judgments. *Social Cognition, 26*, 447–468.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior, 16*, 107–112.

Johnston, W. A., Dark, V. J., & Jacoby, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 11*, 3–11.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition, 8*, 338–342.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9*, 45–48.

Schwartz, M. (1982). Repetition and rated truth value of statements. *American Journal of Psychology, 95*, 393–407.

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 219–230.

Westerman, D. (2008). Relative fluency and illusions of recognition memory. *Psychonomic Bulletin & Review, 15*, 1196–1200.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1235–1253.

Whittlesea, B. W. A., & Leboe, J. P. (2003). Two fluency heuristics (and how to tell them apart). *Journal of Memory, and Language, 49*, 62–79.

Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica, 98*, 141–165.

Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 547–565.

Whittlesea, B. W. A., & Williams, L. D. (2001a). The discrepancy-attribution hypothesis: I. The heuristic basis of feeling of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 3–13.

Whittlesea, B. W. A., & Williams, L. D. (2001b). The discrepancy-attribution hypothesis: II. Expectation, uncertainty, surprise, and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 14–33.

Willems, S., & Van der Linden, M. (2006). Mere-exposure effect: A consequence of direct and indirect fluency-preference links. *Consciousness and Cognition, 15*, 323–341.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplement, 9*, 1–27.

Zajonc, R. B., Swap, W. C., Harrison, A. A., & Roberts, P. (1971). Limiting conditions of the exposure effect: Satiation and relativity. *Journal of Personality and Social Psychology, 18*, 384–391.