Memory strategies mediate the relationships between memory and judgment

Silvio Aldrovandi (S.Aldrovandi@city.ac.uk)
Marie Poirier (M.Poirier@city.ac.uk)
Department of Psychology, City University, Northampton Square, London, EC1V 0HB, UK

Daniel Heussen (Daniel.Heussen@psy.kuleuven.be)
Department of Psychology, University of Leuven, Tiensestraat 102, 3000 Leuven, Belgium

Peter Ayton (P.Ayton@city.ac.uk)
Department of Psychology, City University, Northampton Square, London, EC1V 0HB, UK

Abstract

In the literature, the nature of the relationships between memory processes and summary evaluations is still a debate. According to some theoretical approaches (e.g., "two-memory hypothesis"; Anderson, 1989) retrospective evaluations are based on the impression formed while attending to the to-be-assessed stimuli (on-line judgment) – no functional dependence between information retrieval and judgment is implied. Conversely, several theories entail that judgment must depend, at least in part, on memory processes (e.g., Dougherty, Gettys, & Ogden, 1999; Schwarz, 1998; Tversky & Kahneman, 1973). The present study contributes to this debate by addressing two important issues. First, it shows how more comprehensive memory measures than those used previously (e.g., Hastie & Park, 1986) are necessary in order to detect a relationship between memory and retrospective evaluations. Secondly, it demonstrates how memory strategies influence the relationship between memory and judgment. Participants recalled lists of words, after having assessed each of them for their pleasantness. Results showed a clear association between memory and judgment, which was mediated by the individual strategies participants used to recall the items.

Keywords: Memory-judgment relationships; On-line judgment; Accessibility

Introduction

How do people provide summary assessments? That is, how do they provide an unitary and coherent judgment about an event that may have changed in quality over time (e.g., Fredrickson & Kahneman, 1993) or about a target stimulus that has been described in a sequential manner (e.g., Lichtenstein & Srull, 1987)?

One appealing and experimentally supported answer to the above question is that people are largely influenced by the memory trace they retain about the to-be-assessed stimuli (e.g., Dougherty et al., 1999; Schwarz, 1998). Biases in retrospective assessments can therefore be comprehended on the basis of a bias in retrieval-based processes (e.g., "Availability"; Tversky & Kahneman, 1993).

However, an alternative series of interpretations have been formulated in the literature where memory does not play a central role (e.g., Anderson, 1989; Betsch, Plessner, Schwieren, & Gütig, 2001). According to these views, people evaluate the stimuli while they experience them (on-line judgment) and rely on this on-line impression in order to provide summary assessments. Retrieving information about the stimuli themselves becomes both cognitively costly and unnecessary in order to provide retrospective evaluations. Fuelling the debate, experimental evidence which failed to obtain correlations between memory and judgment supports this "Independence" view (see Hastie & Park, 1986).

Memory measures and individual strategies

In the present paper, it is argued that a number of issues must be addressed before the memory-judgment relationship can be adequately understood.

Together with several other studies which investigated memory-judgment correlations (e.g., Lichtenstein & Srull, 1987; Reyes, Thompson, Bower, 1980) Hastie and Park (1986) used a ratio measure for the memory output. For instance, in their Experiment 1, participants heard a recorded 5-min conversation between two people, after which they had to assess the job suitability of one of the two characters. The authors then computed a memory ratio by dividing the positive arguments recalled (i.e. those supporting candidate suitability) by the total number of arguments remembered: the higher the ratio, the more favourable the memory for the specific candidate. This memory measure was then correlated with the overall job suitability rating the participant provided for the hypothetical character.

An implicit assumption underlying the use of this type of ratio is that each argument recalled (and possibly each item on participants' mind at the time of judgment) has the same weight in the overall evaluation. Such an assumption is disputable in light of the work reviewed above showing significant biases in the weight various elements have in retrospective judgments. In addition, the availability heuristic (Schwarz, 1998; Tversky & Kahneman, 1973) suggests that the ease of recall of a given piece of information mediates its effect on evaluations. Accordingly, it can be argued that the availability of each item at the time of judgment will impact on retrospective evaluations: the more easily accessible a specific item is at the time of

evaluation, the higher its influence on the evaluation itself. This would in turn entail that a memory measure (like a ratio) which assigns the same weight to each item recalled does not adequately represent the memory content accessed in order to produce overall assessments.

Second, it is argued that recall strategies can severely impact the relationships between memory and judgment. In most research which investigated memory-judgment correlations, the memory task usually follows the judgment task (see Hastie & Park, 1986). The nature of the stimuli representation accessed at the time the memory task is prompted can therefore be different from the representation of the stimuli accessed at the time an overall assessment is produced. One possible mechanism which can increase the difference between these two representations is the recall strategies participants adopt. In the present paper we focus on serial recall strategies, i.e. the tendency to recall the items from the just-presented series in the order in which they were presented. Relying extensively on this memory strategy could cause the last item, which possibly affected retrospective evaluations due to its high *accessibility*, to have a diminished possibility of being recalled at the later stage. This in turn could hamper the correlations between memory and judgment.

Study

In this study the predictions of the memory-based approach to retrospective assessments are tested through various means. First, correlations between memory and judgment were examined. In doing so, we compared different measures. We computed a "global", ratio-type, memory measure, as this was called upon in previous studies and we have argued that it may have masked the relationship between memory and judgment (e.g., Hastie & Park, 1986; Lichtenstein & Srull, 1987). This memory ratio measure reflected the degree of isolation with which a distinctive (negative) item was recalled: the higher the ratio, the fewer neutral items were recalled together with the negative item. We then correlated this measure of memory with a retrospective judgment measure.

Second, judgment was required first and memory for the content of the word-list obtained second (details of how this was done follow below). Hence, it was possible to contrast the mean pleasantness rating obtained when the negative item was recalled with the mean rating for the trials when it was not recalled. Presumably, if a distinctive item is available for later recall, it is more likely to have been available at the time of judgment; conversely, if the negative item is not available for recall, the probability that it was available at the time of judgment is reduced. Hence, we would expect that on average, the pleasantness rating will be lower in the cases where the negative item was available for the memory component of the task.

Third, the mediating role of recall output strategies on the memory-judgment correlations was investigated. Participants were administered a "free" recall task, where they were provided no constraints about the order in which they could recall the items[1]. In doing so, it was possible to investigate participants' preferential output strategies by comparing the items presentation order with the item recall order; participants were then divided into two groups, depending on the extent to which they recalled the items in the order in which they were presented to them. It is predicted that different memory-judgment correlations pattern will arise between the two groups: Participants who exhibit serial recall strategy to a larger extent in their response will produce diminished memory-judgment correlations for items presented in later positions.

Last, the effect of negative item availability was further analysed by examining ease of recall. As a measure of relative memory accessibility, we used output position in the memory task. Since participants were asked to perform free recall (hence no output constraints were implemented), we made the simplifying assumption that items recalled first are on average more readily accessible in memory. It was assumed that negative items recalled early on were more easily accessible than negative items that are recalled later on and would have had more impact on the retrospective assessment. Hence, our hypothesis was that the earlier a negative item was recalled, the stronger its impact on retrospective evaluations. For this reason, the effect of this differential accessibility of negative words was analysed in relation to the retrospective evaluations participants provide for the *lists as a whole*.

Method

Participants Thirty-six undergraduate students (31 females) from City University London took part in the study. Age ranged from 18 to 56 years (M = 26.7, SD = 10.6). Participants were granted course credits for an introductory course in psychology for their participation.

Design and Materials A pool of 192 words was selected from the Affective Norms of English Words database (ANEW; Bradley & Lang, 1999). Twenty-four negative items were selected along with 168 neutral ones. The selection was based on the valence (i.e. how positive or negative they are) and arousal scores (e.g., how much activation the normative sample reported on reading the word) of each item on database scales. Negative items were selected to be low in valence (less than 3, on a scale of 1-9) and high in arousal (over 5.9, on the same scale). Neutral items scored in the middle range for valence (4.5 to 6.9) and low on the arousal scale (less than 5). The selected negative items scored significantly lower than neutral items on valence scores (M = 2.3 and M = 5.7, respectively), $t(190)$ = 25.3, $p < .001$ and had a significantly higher arousal rating (M = 6.5 and M = 4.1, respectively), $t(190)$ = 21.1, $p < .001$.

[1] A pilot study revealed that the combination of a judgment task with a serial recall task is too demanding for participants to pay enough attention to both tasks. Participants tended to perform the judgment task in a cursory manner.

From the resulting word pool, 32 six-word lists were created, as follows. Eight lists included a negative item in the first position followed by 5 neutral words—hereafter identified as "Start" lists. Eight "Middle" lists had a negative item in the middle positions (4 lists in 3rd position and 4 lists in 4th position.)[2] Eight "End" lists comprised five neutral items and a negative word in last position. Finally, eight control lists contained only neutral words.

Within-list matching between the negative (if any) and the neutral items ensured that negative and neutral words were equated on familiarity ratings (Coltheart, 1981), number of phonemes, and the Kucera-Francis frequency index (Kucera & Francis, 1967).

Procedure All participants took part in individual testing sessions that lasted approximately 45 minutes. The experiment was controlled by a computer program developed specifically for the present experiment with Authorware 7.0 (Adobe / Macromedia, 1987, 2003). A series of introductory screens familiarised participants with the computer-controlled procedure and gathered demographic data.

Participants were told that the aim of the experiment was to collect normative data about the pleasantness of 6-word lists. They were instructed to attend to the lists and to provide an overall pleasantness rating for each one immediately after its presentation. The ratings were on a 0-100 scale (0 = *very unpleasant*, 100 = *very pleasant*), and participants were encouraged to make use of the whole range in their responses.

Each word was presented for one second with an inter-stimulus interval of 0.75 seconds. A series of asterisks appeared on the screen for 3 seconds to signal the end of the list presentation. After the asterisks had disappeared from the screen, participants were prompted to provide their rating – for which they had no time limit. Participants were required to use the mouse to click on a slide bar (with extremes of 0 and 100) on the position they felt was closest to their impression of the list. In order to limit the extent of anchoring effects (e.g., Chapman & Johnson, 2002) a sliding marker would appear on the bar (with its equivalent numerical value underneath) only after participants clicked for the first time on the slide bar. Participants then had the opportunity to adjust this initial rating by sliding the marker, and were to confirm their final one by clicking on a "Continue" button.

After rating a given list's overall pleasantness, participants were required to perform a free recall task, i.e. they were asked to type all the words they could remember from the most recently presented list. The instructions emphasised that spelling errors would not affect scoring and that both the assessment and recall tasks were equally important; participants were asked not to overlook the rating task in order to proceed more quickly to the recall task. After participants had completed the recall task, the next list was presented, and so on.

Three practice trials were provided. List presentation order was randomised independently for each participant and no time limit was set for either the rating or the recall tasks. A post-experimental questionnaire was used at the end of the session in order to gather information about how participants completed both the judgment and recall tasks.

Results

Due to the dual nature of the task, a precautionary measure was taken in order to exclude from the analyses any participant who neglected the judgment task in order to proceed more quickly to the memory task. Participants whose judgment scores were characterised by a standard deviation of 5 or less (5% of the scale) were eliminated from the analyses. One participant was excluded according to this criterion. Moreover, in the post-experimental questionnaire, the same participant indicated that s/he performed the pleasantness ratings according to list memorability rather than perceived pleasantness. Alpha was set to .05 for all analyses.

Judgment Figure 1 presents the mean pleasantness ratings as a function of list type. Ratings for the Control lists were the highest (M = 60.8, SD = 12.4). Moreover, lists with a negative item in either first (Start lists; M = 44.3, SD = 13.3) or last position (End lists; M = 44.2, SD = 12.5) were rated as more unpleasant than Middle lists (M = 49.7, SD = 12.0).

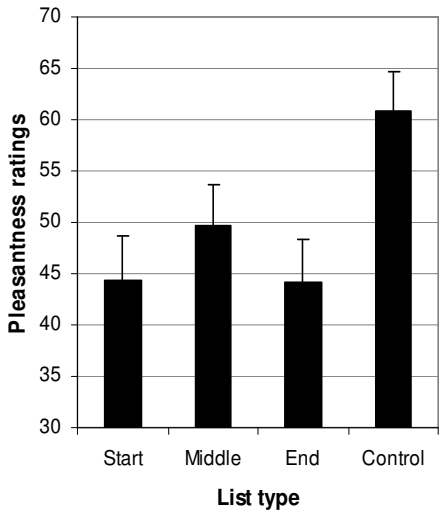

Figure 1: Mean pleasantness ratings as a function of list type. Error bars represent 95% confidence intervals.

A one-way ANOVA was run, with list type (Start, Middle, End and Control) as the within-subject factor. Overall, there was a significant main effect of list type, $F(3, 102) = 52.5$, $p < .001$, $\eta_p^2 = .61$. Planned contrasts indicated that

[2] Analyses revealed no differences in either memory or judgment measures between lists with a negative item in 3rd or 4th position.

pleasantness ratings for Start, Middle and End lists were significantly lower than for Control lists (all *p*s < .001). Planned contrasts also revealed that the pleasantness ratings for Start (primacy) lists were lower than for Middle lists, $t(34) = 4.8$, $p < .001$, $d = .81$. Moreover, a recency effect was observed too as the ratings for End lists were lower than for Middle lists, $t(34) = 5.2$, $p < .001$, $d = .88$.

Memory Figure 2 represents the mean recall proportion for the negative item as a function of word position and valence. It seems that overall there was a memory advantage for the negative items as compared to the neutral ones presented in the same position; participants exhibited higher recall rates for the negative items than for the neutral ones, across presentation positions. More importantly, memory primacy and recency effects for the negative items can be observed: Negative items presented in either first ($M = .78$, $SD = .17$) or last position ($M = .71$, $SD = .23$) were better recalled than those presented in the middle positions ($M = .56$, $SD = .23$).

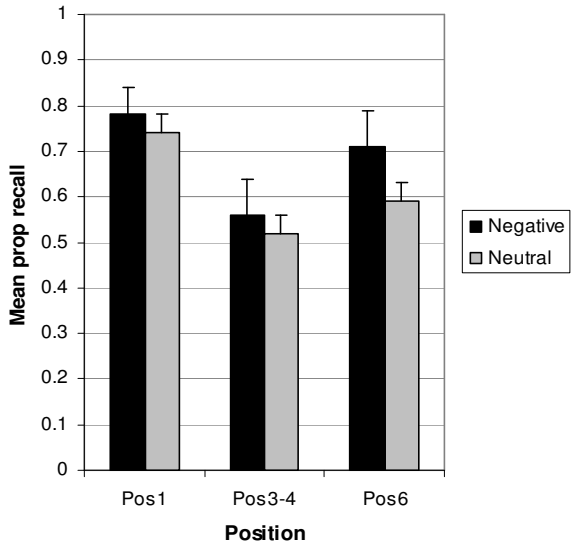

Figure 2: Mean proportion recall as a function of word position and valence. Error bars represent 95% confidence intervals.

The data were analysed using a 2 (valence: negative vs. neutral) × 3 (position: 1st, 3rd/4th, and 6th) repeated measure ANOVA. Main effects of position ($F(2, 68) = 26.9$, $p < .001$, $\eta_p^2 = .44$) and valence ($F(1, 34) = 12.0$, $p < .01$, $\eta_p^2 = .26$) were noted. The valence by position interaction was not significant ($F(2, 68) = 1.4$, $p = .25$), indicating that the memory advantage for the negative items over the corresponding neutrals was relatively constant across positions.

To test for primacy and recency effects in the recall of the negative items, planned contrasts analyses were computed. The results showed that negative items presented in the middle positions were recalled less than those presented in first ($t(34) = 5.5$, $p < .001$, $d = .92$) and last positions ($t(34) = 2.9$, $p < .01$, $d = .49$). Hence, primacy and recency effects were observed for the recall of the negative items.

Memory-Judgment relationships In order to reduce the influence of potential anchoring effects (e.g., Chapman & Johnson, 2002) and of inter-individual differences in the use of the 0-100 scale, judgment scores were transformed as follows: For each participant, the average pleasantness rating for the Control lists was subtracted from the pleasantness ratings for each Start, Middle and End list – that is the lists that contained a negative item. The new corrected judgment scores (J') therefore represented how much more unpleasant each Start, Middle and End list was in comparison to the average Control list for each participant. J' scores were then averaged for each participant, according to the negative item presentation position and whether the negative item presented in the list was recalled or not.[3]

1. "Global" correlation First, a "global" memory score for each participant was computed (e.g., Hastie & Park, 1986). This score was the average ratio between the negative information recalled and the total amount of words recalled: the higher the value, the more *negative* the memory for the list. Thus, the lowest possible score of 0 corresponds to neutral items only being recalled and the maximum score of 1 refers to lists where only the negative item was recalled. We then correlated this measure with the overall average corrected judgment (J') for each participant: It was expected that the more negative the memory for the list, the lower the pleasantness ratings. However, the correlation yielded non significant results, $r(35) = -.133$, $p = .45$, and revealed how memory and judgment measures were not associated – at least when using this ratio-style memory measure.

2. Judgment according to negative item recall Second, we computed more comprehensive memory measures: we compared the corrected average pleasantness rating for lists where the negative item was recalled versus lists where the negative item was not recalled. Overall, when the negative item was recalled in the memory task, pleasantness ratings were lower ($M = -16.7$, $SD = 11.4$) than when the negative item was not recalled ($M = -9.9$, $SD = 9.5$). When these results were broken down by list-type, the same pattern appeared for Start and Middle lists. For End lists, the pleasantness ratings were low regardless of whether the negative item was recalled or not.

A 2 (Memory: Negative item recalled Vs. not recalled) × 3 (List type: Start, Middle and End) within-subjects ANOVA confirmed these observations. The main effect of Memory was significant ($F(1, 34) = 20.4$, $p < .001$, $\eta_p^2 =$

[3] This analysis yielded a total of 6.2% missing values. Missing values were replaced using different methods, including mean substitution by subject, grand mean, and Expectation-Maximization algorithm (Schafer & Olsen, 1998). As all the analyses returned the same results, we will be reporting the data obtained via mean by subject substitution.

.38), confirming that overall ratings were more unpleasant for those lists where the negative word was recalled. The List by Memory interaction was significant, too ($F(2, 68) = 5.0$, $p < .01$, $\eta_p^2 = .13$). Follow-up analyses revealed that the main effect of Memory was significant for Start and Middle lists ($t(34) = 4.3$, $p < .001$, $d = .74$; $t(34) = 3.0$ $p < .01$, $d = .51$, respectively). However, for End lists the pleasantness ratings were invariably low regardless of the negative item being recalled or not ($t(34) = 1.2$, $p = .24$). Table 1 below summarises these findings.

Table 1: Mean corrected pleasantness ratings (J') as a function of list type and negative item being recalled or not

		List Type		
		Start	Middle	End
Was the negative item recalled?				
No	*M*	-7.7	-7.1	-14.9
	SD	(12.9)	(10.5)	(14.8)
Yes	*M*	-19.1	-13.8	-17.2
	SD	(13.3)	(12.5)	(11.0)

3. The role of memory strategies In order to investigate why no relationship between memory and judgment was observed for End lists, the extent to which participants serially recalled the items was investigated. The rationale behind this analysis goes as follow: As soon as the negative item presented in last position disappeared from the screen, its availability in memory was still high and hence it influenced the retrospective judgment; however, if a participant tended to output the items in the order in which they were presented to them, the likelihood of that negative item to be produced at the recall stage diminished (e.g., because of output interference; Nairne, 1990).

In order to test this hypothesis, participants were divided into 2 groups, depending on their recall output strategies. Fourteen "Serial recallers" (SR) obtained the top 40% score in strict serial recall; on the other hand, "Non serial recallers" (NSR; n = 13), scored at the bottom 37%.[4] Table 2 below shows that SR achieved considerably higher strict serial recall scores across positions ($F(1, 25) = 31.1$, $p < .001$, $\eta_p^2 = .56$).

Table 2: Mean proportion strict serial recall depending on item presentation position for Serial (SR) and Non serial recallers (NSR)

	Pos1	Pos2	Pos3	Pos4	Pos5	Pos6
SR	.67	.48	.36	.25	.16	.14
NSR	.34	.17	.14	.06	.02	.01

[4] The middle 23% of the sample was excluded from the analyses in order to increase the distinction between the two groups. However, the results did not substantially change when the allocation to groups was median-based.

Two separate 2 × 3 within-subjects ANOVAs were ran[5], with Memory (Negative item recalled Vs. not recalled) and List type (Start, Middle and End) as the factors. The analyses revealed how for NSR the main effect of Memory was significant ($F(1, 12) = 11.2$, $p < .01$, $\eta_p^2 = .48$) – but the interaction List type by Memory was not ($F < 1$). Taken altogether these analyses confirmed that overall ratings were more unpleasant for those lists where the negative word was recalled – and this was true regardless of the presentation position of the negative item. Conversely, for SR, both the main effect of Memory ($F(1, 13) = 9.0$, $p < .05$, $\eta_p^2 = .41$) and the interaction Memory by List Type ($F(2, 26) = 5.1$, $p < .05$, $\eta_p^2 = .28$) were significant. Follow-up analyses revealed how the association between memory and judgment measures was observed for Start and Middle lists ($t(13) = 3.7$, $p < .001$, $d = .99$; $t(13) = 2.5$ $p < .01$, $d = .66$, respectively); on the other hand, for End lists the pleasantness ratings were comparably low regardless of the negative item being recalled or not ($t(13) = 0.1$, $p = .90$)[6].

4. Accessibility Finally, retrospective evaluations were analysed depending on the negative item recall position (see Table 3). In other words, the pleasantness ratings (J') were examined according to the position in which the negative item was recalled by the participants (regardless of its presentation position). The underlying rationale was that items that are more accessible in memory are likely to be recalled earlier—if the negative item is more accessible and recalled early we would expect its impact on retrospective evaluations to be higher than when it is recalled later in the protocol or not at all.

Table 3: Mean corrected pleasantness ratings (J') as a function of negative item recall output position

		Negative item recall output position		
		Not recalled	Positions 1 & 2	Positions 3 to 6
Pleasantness ratings (J')	*M*	-10.3	-18.9	-13.8
	SD	(10.7)	(11.7)	(10.8)

Overall, the pleasantness ratings varied depending on the negative item output position: They were lowest when the participants recalled the negative item as either the first or second response ($M = -18.9$, $SD = 11.7$).

The analysis revealed a significant main effect of recall position on pleasantness ratings ($F(2, 68) = 16.0$, $p < .001$, $\eta_p^2 = .32$). Planned contrasts confirmed that judgments were

[5] These analyses yielded a total of 7.4% missing values – which were once again replaced via mean by subject substitution.

[6] The same conclusions were also reached through analyses where no missing values imputation was required. For instance, the association between memory and judgment was significant for the whole sample also for End lists ($F(1, 28) = 5.2$, $p < .05$) – once the variability due to strict serial recall scores was accounted for ($F(1, 28) = 4.7$, $p < .05$).

lower when the negative item was recalled amongst the first two responses than when it was recalled amongst the last four responses, $t(34) = 4.3$, $p < .001$, $d = .72$, which in turn were lower than for those lists where the negative item was not recalled at all, $t(34) = 2.3$, $p < .05$, $d = .38$.

Discussion

The results of the present experiment largely support the view that memory plays a central role in retrospective evaluations (e.g., Dougherty et al., 1999; Schwarz, 1998).

First of all, both primacy and recency effects were observed for summary evaluations and negative items recall. Negative items presented either at the beginning or at the end of the series exerted a larger impact on pleasantness evaluations – and were better remembered at a later stage.

The results of the correlational analyses produced different outcomes, depending on the memory measure that was being used. When a global, ratio-style, memory measure was utilised (e.g., Hastie & Park, 1986; Lichtenstein & Srull, 1987) – where each recalled item holds the same weight – no significant associations were observed between memory and judgment. However, when the role of memory was tested through more comprehensive measures – or measures which also take item accessibility into consideration (e.g., Schwarz, 1998) – the results provided support for a memory-based approach.

Pleasantness ratings for those lists where the negative item was recalled were significantly lower than for those lists where the negative item was not recalled. This result supports the idea that when a negative item was easily available in memory at the time of judgment it exerted a higher impact on judgment. The assumption was that when a negative item was not recalled in the memory task, it was also less likely to be available at the time of retrospective evaluations; on average, this would lead to a less negative assessment of the list.

These results seem to suggest that participants relied at least to some extent on episodic information stored in memory when providing retrospective evaluations. If they had exclusively relied on on-line judgment formation, there would be no reason to expect an association between memory and judgment measures, although as is always the case with a correlational approach a third factor could perhaps be causing changes in both memory and retrospective evaluations. However, accessibility in memory of a negative item seemed to mediate retrospective judgment, since lower ratings were associated with the negative item being recalled early in the response sequence. The on-line view does not lead to the expectation that the accessibility of the distinctive-negative item would have an impact on retrospective evaluation.

Finally, some of the more detailed follow-up analyses showed how the above mentioned association between memory and judgment was not observed when the negative item was presented in last position. Results suggest that one possible explanation lies in the recall strategies participants used – as clear associations between memory and judgment for End lists were observed for participants who constrained to a less extent the order in which to recall the items.

References

Anderson, N. H. (1989). Functional memory and on-line attribution. In J. N. Bassili (Ed.), *On-line cognition in person perception* (pp. 175–220). Hillsdale, NJ: Erlbaum.

Betsch, T., Plessner, H., Schwieren, C. & Gütig, R. (2001). I like it but I don't know why: A value-account approach to implicit attitude formation. *Personality and Social Psychology Bulletin, 27*, 242–253.

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. *Technical report C-1,* Gainesville, FL. The Center for Research in Psychophysiology, University of Florida.

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgments of belief and value. In T. Gilovich, D. Griffin, & D. Kahneman, (Eds.), *Heuristic and Biases* (pp.120–138). New York: Cambridge University Press.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, 33A*, 497–505.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review, 106,* 180–209.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect and in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology, 65,* 45–55.

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

Kucera, H., & Francis, W. N. (1967). Computational Analysis of Present-Day American English. Brown University Press, Providence R.I.

Lichtenstein, M., & Srull, T. K (1987). Processing objectives as a determinant of the relationship between recall and judgment. *Journal of Experimental Social Psychology, 23,* 93–118.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18,* 251–269.

Reyes, R. M, Thompson, W. C., & Bower, G. H. (1980). Judgmental biases resulting from differing availabilities of arguments. *Journal of Personality and Social Psychology, 39,* 2–12.

Schafer, J. L., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research, 33,* 545–571.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review, 2,* 87–99.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232.