

Supplementary Material

Materials and Methods

Supplementary Table 1. Demographics of the four groups.

	Control Emotional (n=24)	Control Neutral (n=23)	Experimental Emotional (n=27)	Experimental Neutral (n=28)
Age (years)	23.75±5.61	24.57±5.87	25.11±5.04	25±5.81
Educational level (years)	15.46±2.9	15.61±2.45	15.85±2.54	14.86±2.66
Sex (n of females)	12	10	14	14

There were no differences between groups in terms of age [$F(3,98)=.308, p=.820$], educational level [$F(3,98)=.704, p=.552$], nor sex representation [$\chi^2(3)=.389, p=.942$].

Values for age and educational level depict means \pm standard deviation; values for sex depict number of female participants. Participants who did not reach learning criteria on day 1 or day 2 (see below) were not called for the day 3, and were excluded from the analyses (n=10).

Materials

Stories

Each story consisted of 164 words. Thirty-two of those words were defined as memory clues, which were further classified as targets (11 words) and fillers (21 words). Targets were emotionally charged words in the negative story and their neutral counterparts in the neutral story (e.g., «good conditions of living» in the neutral story was substituted for «miserable conditions» in the emotionally charged story). Fillers were emotionally neutral words in both stories.

In summary, the neutral story was about a wealthy family with a daughter studying pharmacist in another city, sending back vitamins to her family, and a belated letter announcing her return and more information about the last box of vitamins she has sent them. The emotionally charged story was about a poor family with a daughter working abroad, sending back vitamins and money to her family, and a belated letter announcing her death and informing that the last box sent to them was her ashes.

Pilot studies

The final versions of the two stories were processed and finalized in the context of two pilot studies. In the two pilot studies, 4 participants -who did not later participate in the main

experiment, were presented with the two original stories, took part in learning (day 1) and recalled the stories two days later (day 2), without interference learning. Based on their emotional ratings, learning and memory performance, as well as comments, the stories were further simplified.

In particular, after pilot 1 (2 participants), very similar words in the original stories (as “food supplements”, “minerals”, “antioxidants”, “vitamins”, and “preparations) were replaced by the word “vitamins”, because their adjacent occurrence was confusing for the participants. After pilot 2 (2 more participants), the sentence “the vitamins had to be kept in air vacuum” in the neutral story was removed, as participants were confused by its meaning. Lastly, as all pilot participants in learning and recall spontaneously replaced the word “broaden” with “expand” (in the sentence “to broaden her horizons”), we substituted the former word with the latter in the final version of the stories.

Images

The two images were selected to be similar to each other (forest scenes), but distinct in emotionality and congruent with story valence. A subgroup of participants in this study ($n=33$) was asked to later rate these images in terms of arousal and valence. In the neutral picture a sunny spring day was presented (for valence: $M = 2.85$, $SE = .22$; for arousal: $M = 1.74$, $SE = .16$), whereas in the emotionally charged picture the landscape was darker, the day was cloudy, and all the plants were dry and lifeless (for valence: $M = 6.27$, $SE = .23$; for arousal: $M = 5.09$, $SE = .34$).

Learning criteria

For learning the stories on day 1 and 2, a maximum of 5 learning trials were administered and each time recalled words were manually recorded by the examiner. A minimum performance criterion entailing correct recall of 80% of the memory clues (25/32 items) was set a priori and participants who did not reach this criterion after 5 learning trials were not called for the next session ($n=10$).

Data coding criteria

Each word or phrase recalled in correspondence to the (32) memory clues of each story was rated in a 4-point scale (with 0.3, 0.5, 0.7, or 1.0). In particular:

- i. A rating of 0.3 was given to recalled items bearing a distant relationship to the context. These cases included: (a) vague definitions (e.g., “she sent some money” instead of “she sent the money from her savings”), (b) the memory clue was distorted but

maintained elements of time, place and context (e.g. “South” instead of “North”), (c) a significant portion of the clue was absent (e.g. “escaped” instead of “escaped from her daily contempt”).

- ii. A rating of 0.5 indicated moderate degree of relationship to the corresponding phrasal memory clue. These included omissions of a single word (e.g., “her family” instead of “her poor family”), substitution of one or more words without significant distortion of the initial context (e.g., “the violence of her family” instead of “the physical violence of her father”), or replacement of a name with one that was not included in the second story (e.g., “Zezet” instead of “Zanet”);
- iii. A rating of 0.7 indicated close relationship to the corresponding memory clue (e.g., “proteins” instead of “vitamins”);
- iv. A rating of 1.0 indicated clues recalled verbatim.

Memory clues recalled in the wrong context (i.e. memory clues from the emotional story recalled as parts of the neutral one, and vice versa) were rated as intrusions.

Supplementary Table 2. Scoring clues of both stories.

Neutral story	Emotional story
1. Relatively wealthy	Very poor*
2. Isabella	Gabriella
3. Miranda	Miriam
4. Janis	Janet
5. Franklyn	Freddy
6. New apartment	Old Shanty*
7. Good living conditions	Miserable living conditions*
8. Isabella the youngest daughter	Gabriella the oldest daughter
9. Raised with daily care	Escaped the daily contempt*
10. Tender love of her mother	Physical violence of her father*
11. She was studying	She was working
12. Miles away	Kilometers away
13. South	North
14. California	Mexico
15. For education	For survival
16. To expand her cognitive horizons	To help her poor family
17. University	Factory
18. She was studying as a Pharmacist	She was working as a cleaner
19. Ideal conditions	Terrible conditions*
20. 65 types	85 types

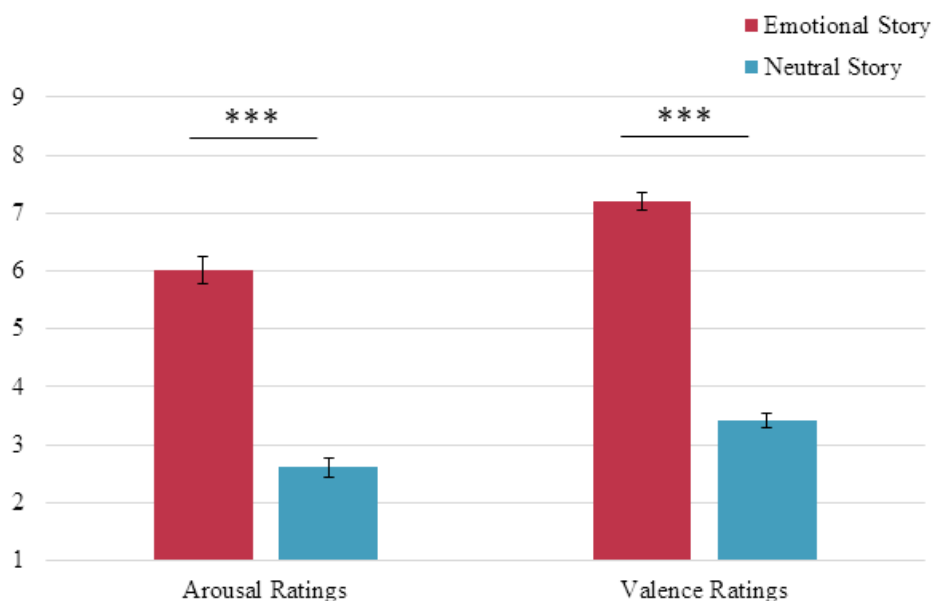
21. Vitamins	Vitamins
22. Antioxidants to spare	Money to spare
23. Plastic box	Glass container
24. Reddish powder	Gray powder*
25. They gave to the oldest daughter	They gave to the youngest daughter
26. One tablespoon a day	Two table spoons a day
27. The arrival of the youngest daughter	The death of the oldest daughter*
28. According to the new studies	According to her last wish*
29. University	Gabriela
30. The plastic box that was sent	The glass container that was sent
31. Contained the vitamins	Contained the ashes*
32. From their best batch	From her burned body*

*Target clues

Results

Arousal and valence ratings

The emotional story was rated as more negative [$M_{emo} = 7.20$, $SE = .15$ vs. $M_{neut} = 3.39$, $SE = .11$, $F(1,98) = 385.59$, $p < .001$, $\eta^2 = .797$] see Supplementary Figure 1]. There was no main effect nor interaction with group in valence ratings [main effect of group: $F(3,98) = 2.596$, $p = .057$; emotion by group interaction: $F(3,98) = .868$, $p = .460$].



Supplementary Figure 1. Participants rated the emotional story as more arousing and negative than the neutral story. Please note that the valence scale in this experiment was presented in the reverse order to that of the original scale. Error bars indicated SEM; *** $p < .001$.

Memory performance as a function of interference

Recall of target clues. In addition to the significant two-way interaction on the retention of target clues, the two main effects were also significant. Thus, after controlling for individual variability in arousal ratings between the two stories, overall retention rates for target memory clues for the emotional story were higher than for the neutral story across groups [$M_{emo} = .89$, $SE = .02$ vs. $M_{neut} = .77$, $SE = .02$; main effect of emotion $F(1,97) = 12.269$, $p = .001$, $\eta^2 = .112$]. Retention rates for target memory clues were only marginally higher for the control groups [$M_{control} = .86$, $SE = .02$ vs. $M_{experimental} = .80$, $SE = .02$; main effect of interference $F(1,97) = 3.378$, $p = .069$]. Moreover, the arousal difference score was significantly related to the target memory clue retention rate [$F(1,97) = 17.398$, $p < .001$, $\eta^2 = .152$].

In the same model, without taking into account arousal as covariate, there was a main effect of emotion [$F(1,98) = 8.945$, $p = .004$, $\eta^2 = .084$] and a main effect of interference [$F(1,98) = 5.741$, $p = .018$, $\eta^2 = .055$] on the retention of target clues, but not an emotion by interference interaction [$F(1,98) = 2.655$, $p = .106$]. Post hoc pairwise tests revealed that group differences persisted without controlling for differences in arousal ratings ($M_{NeutExp} = .71$, $SE = .04$ vs. $M_{EmoExp} = .87$, $SE = .04$; $p < .01$).

In additional analysis including difference in valence ratings between the two stories as a covariate, no relation was found between the difference in valence ratings and the retention rates for target memory clues [$F(1,96) = .059$, $p = .808$].

Recall of filler clues and intrusions. There was no main effect nor interaction with emotion, but a main effect of interference on the retention of filler clues. Specifically, after controlling for individual variability in arousal ratings between the two stories, overall retention rates for filler memory clues did not differ between the emotional and the neutral story across groups [$M_{emo} = .81$, $SE = .02$ vs. $M_{neut} = .79$, $SE = .02$; main effect of emotion $F(1,97) = .763$, $p = .385$]. In addition, retention rates for filler memory clues were significantly higher for the control groups [$M_{control} = .86$, $SE = .02$ vs. $M_{experimental} = .74$, $SE = .02$; main effect of interference $F(1,97) = 16.952$, $p < .001$, $\eta^2 = .149$]. Lastly, the arousal difference score was significantly related to the filler memory clue retention rate [$F(1,97) = 9.339$, $p = .003$, $\eta^2 = .088$]. In additional analysis including difference in valence ratings between the two stories as a covariate, no relation was found between the difference in valence ratings and the retention rates for filler memory clues [$F(1,96) = 1.715$, $p = .368$].

While there were no differences between emotional and neutral stories in the number of intrusions recalled on day 3 ($M_{emo} = 2.58$, $SE = .46$ vs. $M_{neut} = 3.27$, $SE = .45$), the arousal difference score was significantly related to the intrusions recalled [$F(1,52) = 6.556$, $p = .013$,

$\eta^2=.112$]. In additional analysis including difference in valence ratings between the two stories as a covariate, no relation was found between the difference in valence ratings and the retention rates for intrusions [$F(1,51)=.189, p=.665$].

Memory performance as a function of perceived story valence

Experimental-neutral group. The negative associations between neutral story retention rates and valence ratings of either the neutral story or the interfering emotional story were weak and failed to reach significance (for full results, see Supplementary Table 2).

Experimental-emotional group. Individual retention rates of the emotional story were significantly and inversely related with individual valence ratings of that story ($r = -.561, p = .002$).

Supplementary Table 2. Correlations between target clue retention rates, arousal and valence ratings in the two experimental groups.

Retention rates for target clues	Experimental Neutral group	Experimental Emotional group
Arousal ratings (neutral story)	.532 ($p=.004$)	.184 ($p=.357$)
Arousal ratings (emotional story)	-.504 ($p=.006$)	-.338 ($p=.085$)
Valence ratings (neutral story)	-.100 ($p=.612$)	-.098 ($p=.626$)
Valence ratings (emotional story)	-.279 ($p=.150$)	-.561 ($p=.002$)

Significant relations (at Bonferroni-corrected $\alpha=.0063$) are depicted in bold.