The Role of Emotion in Multimodal Integration

Rémy Versace and Marylène ROSE

Electronic version
URL: http://journals.openedition.org/cpl/1402
ISSN: 1379-6100

Publisher
Centre PsyCLÉ

Electronic reference

This text was automatically generated on 9 May 2019.

© All rights reserved
The Role of Emotion in Multimodal Integration

Rémy Versace and Marylène ROSE

Many studies going back many years have shown that emotion has an effect on memory performance (e.g., Blaney, 1986; Bower, 1981; Parrott & Spackman, 2000). Bower (for example 1981, 1994) initially investigated the beneficial effect of an emotional context on memorization and proposed an explanation which linked this effect with the attentional mechanisms: information with a high emotional content captures the attention more and is therefore remembered better than non-emotional information (see also, Christianson & Loftus, 1991; Christianson, Loftus, Hoffman & Loftus, 1991; Hamann, Ely, Grafton & Kilts, 1999).

Bower also examined the phenomena known as “mood congruity effect” and “mood-state dependent retrieval”. The first refers to the tendency of individuals to retrieve information more easily when it has the same emotional content as their current emotional state. This has been demonstrated both for explicit retrieval (for a review, see Bower, 1981), as well as for implicit retrieval (Watkins, Vache, Vernay & Muller, 1996). "State dependence" refers to the fact that the retrieval of information is more effective when the emotional state at the time of retrieval is similar to the emotional state at the time of memorization. These two phenomena, the mood congruity effect and mood-state dependent retrieval, are ultimately fairly similar to the context effects which have been traditionally observed in memory research (for a review, see, for example Baddeley, 1993; Davis & Thomson, 1988).

While there are numerous arguments to demonstrate the indisputable importance of the attentional and contextual mechanisms, our aim in this research was to show that...
another mechanism can also partly explain the effects of emotion on memorization. This is the integration mechanism. An increasing number of studies in the field of the neurosciences have shown that many different areas of the brain are almost systematically involved in cognitive functioning (e.g., Martin & Chao, 2001; Martin, Haxby, Lalonde, Wiggs & Ungerleider, 1995; Martin, Ungerleider, & Haxby, 2000; Martin, Wiggs, Ungerleider et Haxby, 1996; Ungerleider, 1995). Thus, the integration and/or synchronization of activity in these areas is necessary at a given moment in order to permit the emergence of coherent knowledge and appropriate behaviour, as well as for the long-term conservation of memory traces. The effectiveness of encoding and memory retrieval depends on the level of integration of the various components of the experience within the memory traces (Versace & Nevers, 2003; Versace, Nevers, & Padovan, 2002, Whittlesea, 1987; 1989). Neuro-anatomical observations seem to indicate that specific structures such as the prefrontal cortex and the hippocampus may be involved in the integration mechanism (Bechara, Tranel, Damasio, Adolphs, Rockland, & Damasio, 1995; Bechara, Tranel, Damasio, & Damasio, 1996; Goldman-Rakic, Scalaideh, & Chafee, 2000; Stuss & Alexander, 1999; Ungerleider, 1995). Indeed, the prefrontal cortex occupies a prominent position due to its many connections with several other regions of the brain. Damasio (1995) speaks in terms of an area of convergence. According to this author, the emergence, in response to a given experience, of coherent, unitary knowledge requires the synchronized reactivation of the various components of the experience. This synchronized reactivation demands both the activation of the areas of convergence and the simultaneous activation of backward projections coming from these areas. Obviously, many other cortical (temporal, parietal and frontal regions) and subcortical (e.g. the thalamus) structures play an important role in multisensory integration, depending on the task at hand.

Thus, any factor that facilitates or disrupts the integration mechanism is very likely to facilitate or disrupt encoding and/or memory retrieval. Many studies have shown that the prefrontal cortex and the amygdalia, already thought to play a part in integration, may also play an important role at the level of emotion (e.g., Bechara, Damasio, Damasio & Anderson, 1994; Bechara et al., 1996; Damasio, 1995; 1997; Davidson & Irwin, 1999; Rolls, 1999). Another structure, the amygdala, also seems to be a good candidate for the association of an emotional state with the components of experiences which are primarily of a sensory-motor nature (Adolphs, Tranel, Damasio, & Damasio, 1994; Adolphs, Tranel, & Damasio, 1998; Bechara, Tranel, Damasio, Adolphs, Rockland & Damasio, 1995; Cahill, Babinsky, Markowitsch & McGaugh, 1995; Davidson, 1998; Gloor, 1990; Lane et al., 1997; Packard & Cahill, 2001; Scott et al., 1997). According to McGaugh (2000; 2002; McGaugh, Cahill, & Roozendaal, 1996), the amygdala plays an important role in the consolidation of memory through its projections, which are mostly reciprocal, to many other areas of the brain (the sensory areas, thalamus, hippocampus, prefrontal cortex, etc.).

The hypothesis tested in this research was therefore that emotion may facilitate the establishment of relations between the multiple components of an experience and thus render the trace more integrated and unitary. At the behavioral level, to our knowledge, it was the first study that directly tested this hypothesis. We used pictures of objects and animals and sounds which are generally associated with these objects and animals. During the first phase, the participants had to judge the level of association between the pictures and the simultaneously presented sounds. Each picture/sound pair was preceded
by a picture with negative valence or a neutral picture. In the second phase, the same pictures of objects or animals were presented to the participants. However, in this case they were either associated with the same sound as in the initial phase (sound congruent with picture) or with a sound different from that used in the initial phase (non-congruent sound). In this second phase, the picture/sound pairs were always presented alone (without emotional induction). The participants had to indicate whether the pictures and sounds were congruent or non-congruent.

Our hypothesis was that emotional activation (in this case, negative) should strengthen the association between the pictures and the sounds during the initial phase. Thus when the pictures are presented with the same sounds in the first and second phases, the judgment of congruence should be faster for the pairs encoded in the negative valence condition than for those encoded in the neutral condition. In contrast, when the pictures are presented with a different sound in the two phases, the judgment of congruence should be faster for the pairs encoded in the neutral condition than in the negative valence condition.

Participants. 40 students from the University of Lyon 2, France, were tested. All had normal or corrected-to-normal vision. No participant was familiar with the issues which were being investigated in this study.

The experiment was carried out on a Macintosh G4 micro computer, with a 17” monitor (Pronitron 17/500), using Psyscope software (Cohen, McWhinney, Flatt, & Provost, 1993). The experimental stimuli consisted of 96 pictures and 72 sounds. 48 pictures were coloured photographs of natural scenes that were used to activate either negative emotions (24 pictures) or neutral states (24 pictures) in the participants. These 48 pictures were selected from a data base of 161 pictures tested by Versace, Augé, Thomas Anterior, and Laurent (2002). The other 48 pictures were coloured photographs of objects (24 pictures) and animals (24 pictures). All the pictures subtended a visual angle of 9.4° vertically and 12.5° horizontally.

The 72 sounds were the sounds of objects or animal calls. 48 of these corresponded to the selected pictures (24 to the objects and 24 to the animals), while the other 24 corresponded to the sounds made by objects (12) or animals (12) other than those selected. All the sounds lasted for 1000 ms and were of equivalent intensity. An additional set of 8 pictures (2 negative and 2 positive natural scenes, 2 objects and 2 animals) and 4 sounds (2 sounds of objects and 2 animal calls) was selected to constitute the practice trials.

Each participant was tested individually in a session lasting approximately 15 minutes. At the beginning of the session, the subjects were seated in front of the microcomputer and were asked to rest their heads on a chin rest positioned 50 cm from the monitor. The experiment was divided into two phases, an encoding phase and a test phase. In the encoding phase, the participants were told they would be presented with pairs of stimuli consisting of one picture and one sound and that their task would be to rate the congruence between the two stimuli. Although the sound and the picture were always congruent, the instructions explained that the participants had to judge whether the sound was more or less similar to what they would have imagined on looking at the picture. They had to respond by positioning the cursor of the mouse on a horizontal scale displayed below the picture at the bottom of the screen. The left and right ends of the scale corresponded respectively to “0” (the sound is not at all what I would have imagined for this picture) and “10” (the sound corresponds exactly to what I would have imagined...
for this picture). Each picture-sound pair was preceded by a neutral or a negative picture. No response was required for these initial pictures. The order of presentation of the objects and animals, and of the neutral and negative pictures was randomized. The pictures presented in the neutral condition for half of the subjects were presented in the negative condition for the other half of the subjects. The 48 experimental trials were preceded by 4 practice trials. Each trial began with a fixation point displayed for 1500 ms in the centre of the screen. A neutral or negative picture was then presented for 3000 ms and was followed by the picture of an object or an animal presented simultaneously with its corresponding sound. The picture remained on the screen until the participant responded.

The encoding phase was immediately followed by the test phase in which the participants were presented with 48 picture-sound pairs. Unlike in the encoding phase, the picture-sound pairs were not preceded by negative or neutral pictures. The 48 pictures were the same as in the encoding phase. However, 24 of them were associated with same sound (congruent pairs), and 24 were associated with a different sound (non-congruent pairs). The participants had to decide as quickly and accurately as possible whether the picture and the sound were or were not congruent by pressing the appropriate key on a button-box with the first fingers of their right and left hands. For half of the congruent pairs and the non-congruent pairs, the picture was in the neutral condition in the encoding phase, and for the other half, the picture was in the negative condition. The order of the different experimental conditions was randomized. A picture associated with a congruent sound for half of the subjects, was associated with a non-congruent sound for the other half of the subjects. The 48 experimental trials were preceded by 4 practice trials. Each trial began with a fixation point displayed for 1500 ms in the centre of the screen. The fixation point was then followed by a picture of an object or an animal presented simultaneously with a sound. The picture remained on the screen until the participant responded.

Mean correct responses latencies and error rates were calculated across subject for each experimental condition. Latencies beyond 1500 ms were removed (less than 3% of the data). Mean correct latencies and error rates for the different experimental conditions are presented in table 1. These data were subjected to repeated measures analyses of variance with valence (negative vs. neutral) and picture-sound relation (congruent vs. non-congruent) as the within-subject variables.

Table 1. Mean Response Times (RTs; in milliseconds) and mean percentages of correct responses in each experimental condition (Standard errors are in parentheses).

<table>
<thead>
<tr>
<th>Picture-sound relation</th>
<th>Encoding Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neutral</td>
</tr>
<tr>
<td>Congruent</td>
<td>RT(ms)</td>
</tr>
<tr>
<td></td>
<td>903 (21)</td>
</tr>
<tr>
<td>Non congruent</td>
<td>899 (18)</td>
</tr>
</tbody>
</table>

The Role of Emotion in Multimodal Integration
An analysis of the percentage of correct responses revealed only a main effect of the picture-sound relation, $F(1,39) = 12.85; p < 0.001$. The percentage of correct responses was lower for congruent picture-sound pairs (82.7%) than for non-congruent picture-sound pairs (89%). An analysis of latencies revealed a marginally significant main effect of the picture-sound relation, $F(1,39) = 3.47; p = 0.07$, and, more interestingly, a significant interaction between picture-sound relation and valence, $F(1,39) = 7.97; p < 0.01$. This interaction is presented in Figure 1. In accordance with our hypothesis, planned comparisons showed that when pictures were presented with the same sound as in the encoding phase (congruent picture-sound), mean RTs were significantly shorter in the negative (880 ms) than in the neutral (903 ms) encoding condition, $F(1, 39) = 4.31; p < 0.05$. In contrast, when the pictures were presented with a different sound from that used in the encoding phase (non-congruent picture-sound), the mean RTs tended to be significantly shorter in the neutral (899 ms) than in the negative (928 ms) encoding condition, $F(1, 39) = 3.67; p = 0.06$. Another way of explaining this interaction would be to say that for pictures encoded in the negative condition, a change in the sound relative to the encoding phase dramatically increased mean RTs ($F(1, 39) = 11.86; p < 0.01$, unlike pictures encoded in the neutral condition for which a change in the sound had no significant effect on mean RTs ($F(1, 39) < 1$).

The results of this study clearly show that the encoding of the pictures and the sound presented in the first phase was greatly influenced by the previous activation of a negative emotion when compared to a neutral condition. Firstly, when the same picture-sound pairs were presented in the test phase, they were processed more rapidly when they had been encoded in the negative condition than in the neutral condition. This result could not be attributed to an arousal effect that would have resulted in the better encoding of both the picture and the sound when a negative emotion was activated. Indeed, the hypothesis of a general arousal effect would also lead us to predict shorter RTs for old pictures presented with new sounds when these pictures were encoded in the negative condition than when they were encoded in the neutral condition. The results
revealed precisely the opposite effect: when an old picture was associated with a new sound, it tended to be processed more slowly when it had been encoded in the negative condition than in the neutral condition. The observed interaction between encoding condition (negative vs. neutral) and sound type (old/congruent vs. new/non-congruent) shows that the encoding condition affected the strength of the link between the picture and the sound, and not only the strength of the picture's and sound's memory trace. When this link was broken in the test phase, the responses slowed dramatically only in the negative encoding condition.

So far, the effects of emotion on memory have been accounted for either in terms of attentional mechanisms or in terms of contextual effects. As far as the attentional mechanisms are concerned, authors have evoked attentional bias for emotional stimuli (Kulas, Conger, & Smolin, 2003), narrowing of attention (Burke, Heuer, & Reisberg, 1992), resource allocation (Ellis & Ashbrook, 1988; Versace, Monteil, & Mailhot, 1993), or more general arousal effects. Contextual effects (mood congruence and state dependence) are explained in terms of the similarity between the encoding context and the retrieval context on an emotional dimension. These effects can be perfectly integrated into a constructive and functional approach to memory (see, for example, Glenberg, 1997; Goldstone & Barsalou, 1998; Schyns, Goldstone, & Thibaut, 1998; Versace et al., 2002), in which memory traces encoded during an experience and memory traces activated by an experience, both depend on an interaction between the present experience properties and the memory content that reflects past experiences properties. This memory approach also attributes a very important role to the mechanism responsible for the integration of the various different components of an experiment. It therefore also predicts that any factor that facilitates or disrupts the integration mechanism should facilitate or disrupt encoding and/or memory retrieval. We thus know that the neuronal structures that are thought to make possible multimodal integration are also involved in the in the mechanisms that are specific to working memory (in particular the prefrontal cortex and the hippocampal region, see for example Courtney, Petit, Maisog, Ungerleider & Haxby, 1998; Courtney, Ungerleider, Keil, & Haxby, 1997; Ungerleider, 1995). Given that these same structures are also involved in the emotional mechanisms (together with the amygdala), we hypothesize that emotion may also influence the integration mechanism. This hypothesis was confirmed by the results of the experiment.

BIBLIOGRAPHY


NOTES

1. As noticed by an anonymous reviewer, it could be argued that interaction is involved here. Integration should in fact involve the perception of a coherent percept. However, we think that the reinforcement, by emotion, of the link between the picture and the sound is already the mark of an effect of emotion on the integration mechanism, even if this integration does not involve the perception of a coherent percept.

ABSTRACTS

The aim of this research was to show that emotion can facilitate the mechanism responsible for integrating the various components of an experience within the memory traces. Participants initially had to judge the degree of association between pictures of objects or animals and sounds presented simultaneously. Each picture/sound pair was preceded by a negative or neutral image. In the second phase, the same pictures were presented to the participants, either associated with the same sound (sound congruent with picture) or associated with a different sound (non-congruent sound). The participants had to indicate whether the picture and the word were congruent or not. The results confirmed our hypothesis by revealing that when an old picture was presented with the same sound as in the encoding phase, RTs were shorter in the negative than in the neutral encoding condition. On the contrary, when an old picture was associated with a new sound, it tended to be processed more slowly when it had been encoded in the negative condition than in the neutral condition.
L’objectif de cette recherche était de montrer que l’émotion peut faciliter le mécanisme d’intégration des différents composants d’une expérience stockés au sein des traces mnésiques. Les participants devaient tout d’abord juger le degré d’association entre des images d’objets ou d’animaux et des sons présentés simultanément. Chaque paire d’image/son était précédée d’une image négative ou neutre. Dans une seconde phase, les mêmes images d’objets ou d’animaux étaient présentées, soit associées aux mêmes sons (sons congruents aux images) soit associées à des sons différents (sons non congruents). Les participants devaient indiquer si l’image et le son étaient congruents ou non congruents. Les résultats ont confirmé notre hypothèse en montrant que lorsqu’une image ancienne était présentée avec le même son que lors de l’encodage, elle était traitée plus rapidement lorsqu’elle avait été encodée dans la condition négative que dans la condition neutre. Au contraire, lorsqu’une image ancienne était présentée avec un son différent de celui de l’encodage, elle avait tendance à être traitée plus lentement lorsqu’elle avait été encodé dans la condition négative que dans la condition neutre.

INDEX

Mots-clés: émotion, intégration, multimodalité

AUTHORS

RÉMY VERSACE
Université Lyon 2, Institut de Psychologie, Laboratoire d’Etude des Mécanismes Cognitifs (EMC), BRON, FRANCE, Remy.Versace@univ-lyon2.fr

MARYLÈNE ROSE
Université Lyon 2, Institut de Psychologie, Laboratoire d’Etude des Mécanismes Cognitifs (EMC), Bron, France, Marylene.Rose@univ-lyon2.fr