

From Emotion to Memory: An ACT-R view on the Somatic Marker Hypothesis

Andrea Stocco (stocco@units.it)
Danilo Fum (fum@units.it)
Department of Psychology,
University of Trieste, Italy

Introduction

The Somatic Marker Hypothesis (SMH; Damasio, 1994) is probably the most important contemporary theory of emotions. According to the hypothesis, the neurological substrates of the emotions are the perceived immediate bodily reactions to environmental stimuli, which can be sensed through internal representations that are continuously updated in the sensory regions of the brain. These somatic representations are conveyed, through sensory pathways, to a convergence area in the orbitofrontal cortex. Within this region, they are associated with other representations conveying contextual information. In this way, the emotional reactions become somatic markers for the previously encountered stimuli that elicited them.

Once formed, somatic markers may be reactivated when the organism faces situations similar to the ones that induced the markers. The organism is then already pre-alerted and pre-disposed to react properly, and unconsciously biased towards certain behaviors.

The Iowa Gambling Task

Most of the empirical evidence supporting the SMH comes from experiments performed with a paradigm known as the Iowa Gambling Task (hereafter IGT; Bechara, Damasio, Damasio, & Anderson, 1994). This task was developed to capture, within a laboratory situation, some important aspects of real-life decision making: uncertainty about the future, lack of perfect information, and the trade-off between immediate and postponed rewards.

In the Iowa Gambling Task, participants are asked to repeatedly select a card from an array of four decks, labeled *A*, *B*, *C* and *D*. Each selection always results in an immediate positive outcome. Decks *A* and *B* carry bigger wins, while *C* and *D* lead to smaller monetary rewards.

Unpredictably, however, a win may also be immediately followed by a subsequent negative outcome. These penalties are arranged so that selecting from *A* and *B* (“bad decks”) will produce an overall loss of money. Therefore, the advantageous strategy is to select from *C* and *D* (“good decks”), that yield an eventual profit.

Normal participants usually start selecting from the bad decks, but end up performing significantly more selections from the good ones. More interestingly, selections from the bad decks are predicted by greater anticipated increases in the skin conductance response (SCR) than selections from the good ones (Bechara et al., 1996; 1997). Since these reactions appear before participants acquire conscious knowledge of the task (Bechara et al., 1997), they were originally taken

as evidence for an implicit mechanism of somatic markers that was sensing the bad strategy.

Conversely, patients with lesions in the orbitofrontal cortex (OFC) do not show any SCR increase while performing the task, and, correspondingly, they remain stuck to the bad decks, unable to switch to the good ones.

The original interpretation for such results has been questioned by other researchers. The main point of the debate is how exactly emotions affect higher-level cognition. Tomb, Hauser, Deldin & Caramazza (2002) pointed out that SCR responses may be dissociated from bad decks by varying the scheduling of losses. Similarly, Maia & McClelland (2004) showed that good performance in the task is accompanied by explicit knowledge of the underlying structure, casting doubts on the supposedly unconscious knowledge carried by the markers. The striking difference between healthy subjects and patients, however, is harder to frame, since it implies a specific decision-making inability in a category of patients whose cognitive skills are reported to be preserved (e.g., Eslinger & Damasio, 1985).

Recently, Fellows & Farah (2005) have proposed that the cause of patients' impairment may be an inability to acquire new stimulus-reward links once preliminary associations have been learned. They were successful at showing that patients' impairment disappears when no reversal of previous expectation is required.

A computational model

In Fum & Stocco (2004) we proposed a revision of the Somatic Marker Hypothesis that was grounded on a functional integration of emotion and memory. We put forward a model that could replicate the basic experimental results. The core of the model was the ACT-R declarative memory system, rewritten in Lisp and provided with special routines to implement a memory-sampling decision process within the IGT.

Emotion and memory

The main tenet underlying the ACT-R theory is that human cognitions is adaptive, and that the retrieval of information reflects the probability of occurrence of events in the environment (e.g., Anderson & Schooler, 1991).

However, sometimes uncommon events need to be recalled fast and not to be forgotten in spite of their rarity. This is vital when such information is of valuable biological importance. Since larger baseline activation interferes with the learning of new facts, the most rational solution is to have vital information on rare events to be easily recallable by means of large associative strengths with environmental cues. Such strength should reflect the biological value of the information itself. As a result, relevant events may be recalled promptly in the context they are more likely to occur, without cluttering working memory as if they were constantly active.

The representations of the organism inner state provide an effective way of encoding the immediate biological value of an event, which can be also used to evaluate its associated cues. In this sense, Damasio's theory is both attractive and convincing.

In our model, neither internal somatic states nor emotions are modeled directly. Their computational counterpart, however, is their emotional impact, which is calculated for each outcome, stored, and eventually used to reinforce immediate associations between cues and events.

This associative value is added to the interassociative strength S_{ij} between chunks, which is calculated according to the frequency-based bayesian estimates which are described in equations in Anderson & Lebiere (1998).

This additional associative factor is mediated by the orbitofrontal cortex, which is also thought to play a role in the active maintenance of somatic information in working memory. The contribution of the orbitofrontal cortex is expressed through a new parameter, η .

Core features of the model

In ACT-R, when the proper goal is attended, the activation of related chunks is given by the sum of their base-level activation and the spreading component.

$$A_i = B_i + WS_{ji}$$

we simply added a third factor that was proportional to the experienced emotional value of the event encoded in j :

$$A_i = B_i + WS_{ji} + nV_i$$

The term V_i is the emotional appraisal of the fact encoded in chunk i , and is the output of the processing of different subcortical regions—most notably the amygdala and the basal ganglia. These regions are known to be sensitive to the magnitude and frequency of rewards, and anatomically project to the OFC. In case of monetary values, the emotional impact is obviously related to their numerical magnitude, and was calculated as $V_i = \log(i) / \log(\max(i))$.

It may be noted that the two contextual components look similar. Indeed, they both reflect the activity of two prefrontal areas (dorsolateral and ventromedial) and perform similar functions over different contents, following a general rule in the prefrontal cortex (Goldman-Rakic, 1996; Schoenbaum & Setlow, 2001).

Other computational models

Computational approaches to emotion have been attempted several times. Most notably, Rolls (2000) has proposed an autoassociator network model of the role of orbitofrontal cortex in dealing with emotionally-charged information. This approach is functionally very similar to ours. Wagar & Thagard (2004) have put forward another neural model of cognitive-affective integration. It is much detailed in mimicking existing neural circuits, but we regard some of its mappings as questionable.

Within the ACT-R community, Roman Belavkin has previously dealt with this topic (e.g., Belavkin, 2003). Belavkin explicitly linked the role of emotion with goal value (G) and noise in goal activation (τ) in ACT-R. Our approach is rather different, but we certainly share the common view that the main

computational role of emotion is to allow further processing of relevant information, although we prefer to obtain this by means of implicit retrieval of associated declarative information. Furthermore, we also share the view that the basic mechanism is to be recollected within the subsymbolic part of ACT-R, although its effects may be manifest on the symbolic side.

Simulations

Because of the term ηV_i , normal participants are more likely to recollect negative outcomes that followed their own choices. When the η parameter is set to zero, the model mimics the behavior of orbitofrontal patients (see Figure 1). In this damaged version, it is completely attracted by positive outcomes, whose baseline activation shadows the negative drawbacks and hinders the spontaneous process of recalling (and re-experiencing) the aversive results.

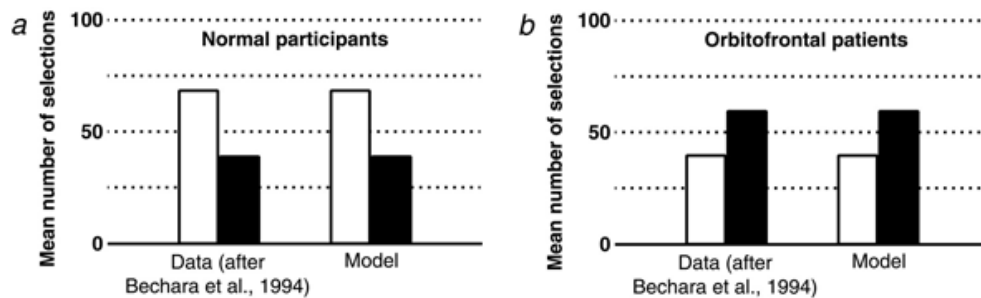


Figure 1: The model reproduces data both from healthy subjects (a) and, when the η parameter is set to zero, of OFC patients (b).

Implicit and explicit processes

Bechara et al. (1997) argued that the effect of somatic markers is entirely implicit: they drive behavior without humans being aware of their action.

We take a different stance. In our model, an automatic and implicit process is required to associate choices with their outcomes, and automatic and implicit is the activation of such information when a particular choice is being attended for evaluation. However, once it has been retrieved, that piece of information is fully explicit and available to conscious processing. This makes possible for a person to select certain options even when they are associated with largest penalties---and larger SCR increases, as in the experiment by Tomb et al. (2002).

More simulations

In addition to its declarative memory store and the orbitofrontal linking mechanism, our model requires other components performing computations. A common way of testing the hypothesized functions of such modules is to disable them and compare our impaired model's performance with that from patients having a functionally corresponding lesion.

The amygdala is known to play a role in the appraisal of frequency and magnitude of rewards (Zalla et al., 2001), and, in particular, to be involved in the processing of fear (LeDoux, 2000). In our model, this immediate appraisal of

outcomes is performed by the function returning the V value. We altered it to return zero for any of the negative outcomes, and were able to obtain a pattern of choices that is similar to what was obtained by Bechara et al. (1999). Results are reported in Figure 2a.

Emotion and working memory in decision making

Bechara et al. (1998) reported an apparent double-dissociation: OFC patients performed normally on working memory tasks but poorly on the IGT; on the contrary, patients with lesions in the dorsolateral part of the prefrontal cortex exhibits severe impairments in working memory but scored normally on the IGT. The authors suggested that decision making may rely on emotional circuits only, and be dissociated from working memory.

Although based on a functional integration of emotion and memory, our model could reproduce this exact pattern of results. A working memory disorder was introduced by reducing the W parameter, and then having the model run the Gambling Task. Our simulated results closely resemble the original data. The rationale underlying our results is that IGT is not an intensive working memory task. A limited amount of attentional resources is required to sample outcomes from memory but, even if this resource is limited, their relative differences on the V value are sufficient to correctly estimate the possible drawbacks from the risky cards.

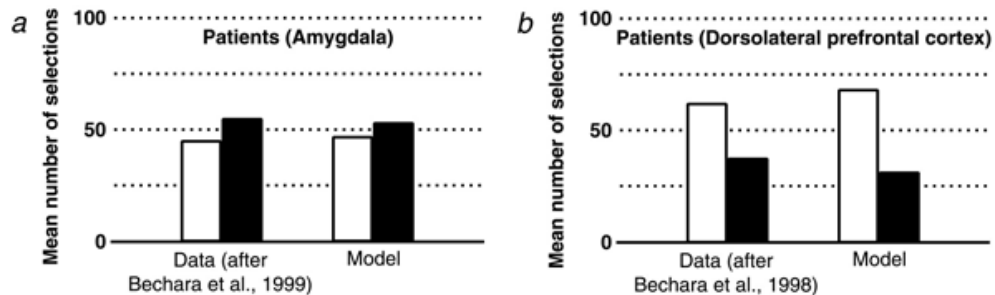


Figure 2. Performance of two damaged versions of the model compared with performance from patients with lesions in the amygdala (a) and in the dorsolateral prefrontal cortex (b).

Conclusions

With our model, we have addressed the issue of the relation between emotion and cognition within the ACT-R approach of the adaptive character of human cognition. We have shown that it can reproduce the basic experimental findings reported by Damasio and co-worker. Finally, we have further test our proposed model, and shown that it can also account for other neuropsychological impairments.

References

- Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396-408.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

- Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15.
- Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, *19*, 5473–5481.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*, 1293–1295.
- Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience*, *18*, 428–437.
- Belavkin, R. V. (2001). The role of emotion in problem solving. In *Proceedings of the AISB'01 symposium on Emotion, Cognition and Affective Computing*, pp. 49-57. Heslington, York, England.
- Damasio, A. R. (1994). *Descartes' Error: emotion, reason, and the human brain*. New York, NY: Gosset/Putnam Press.
- Fellows, L. K., & Farah, M. J. (2005). Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex*, *15*, 58-63.
- Fum, D., & Stocco, A. (2004). Memory, emotion, and rationality: An ACT-R interpretation for Gambling Task results. In C. D. Schunn, M. C. Lovett, C. Lebiere & P. Munro (Eds.), *Proceedings of the Sixth International Conference on Cognitive Modelling*. Mahwah, NJ: Lawrence Erlbaum.
- Maia, T. V., & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences*, *101*, 16075-16080.
- Rolls, E. T. (2000). *The brain and emotion*. Oxford, UK: Oxford University Press.
- Schoenbaum, G., & Setlow, B. (2001). Integrating orbitofrontal cortex into prefrontal theory: Common processing themes across species and subdivisions. *Learning & Memory*, *8*, 134-147.
- Tomb, I., Hauser, M., Deldin, P., & Caramazza, A. (2002). Do somatic markers mediate decisions on the gambling task? *Nature Neuroscience*, *5*, 1103–1104.
- Zalla, T., Koechlin, E., Pietrini, P., Basso, G., Aquino, P., Sirigu, A., & Grafman, J. (2000). Differential amygdala responses to winning and losing: A functional magnetic resonance study in humans. *European Journal of Neuroscience*, *12*, 1764–1770.