Computational approaches to affective processes : evolutionary and neural perspectives

den Dulk, P.

Publication date
2002

Citation for published version (APA):
den Dulk, P. (2002). *Computational approaches to affective processes : evolutionary and neural perspectives*. [Thesis, fully internal, Universiteit van Amsterdam].

Chapter 5

A Connectionist Model for Affective Priming

The dual route model of LeDoux, presented in the previous chapter, may not only function as a model for fear conditioning, but also as a more general view on the relationship between conscious and nonconscious emotion processes. A dissociation between these two processes, similar to that found in conditioning experiments, has been found in the human affective priming research (e.g., Murphy & Zajonc, 1993). In these experiments a sub-optimal prime influenced (i.e., primed) the affective evaluation of a subsequent stimulus, whereas the optimal presentation showed no influence, and even seemed to show a slight (marginally significant) reversal of the effect. Murphy and Zajonc (1993) hypothesized that the effects in the sub-optimal condition were due to a diffuse and nonspecific influence outside of awareness when there is insufficient activation of higher cognitive centers. In the optimal condition, however, higher cognitive processes corrected for the influence of the irrelevant prime.

The architecture of the present connectionist model based on the connectionist model presented in the previous Chapter and on an analysis of the representational requirements of the affective priming task. For the sake of simplicity, elements that were not essential for affective priming, such as learning (i.e., dynamical weight modification), were dropped from the model in the initial simulations. All connections in the model were first set by hand, not by way of parameter fitting, but on the basis of a global analysis. The ensuing model, which was more simple than the fear-conditioning model (e.g., had fewer nodes), revealed differential effects of processing in the direct and indirect routes similar to experimental results. No claim is made that the model even approaches the simulation of consciousness. The differential effects of conscious and nonconscious processing can, however, be obtained even in this relatively simple version of the dual-route model.

Following Jacobs and Nadel (1985), LeDoux (1996) has claimed that stress can bring back responses that were extinguished through the action of the indirect route, but were still preserved in the direct route. Jacobs and Nadel (1985) emphasized the suppression of hippocampal activity, which is presumably responsible for the processing of the context of an event. In humans, a similar phenomenon may be observed when anxiety disorders treated with behavior (extinction) therapy return in stressful situations (Bouton & Swartzentruber, 1991). The affective priming paradigm also provides situations in which the indirect route counteracts processing in the direct route. Shifts in the balance between the functioning of the two pathways are also investigated in the present modeling effort.

The general principle underlying LeDoux' model for affective processing (i.e., fast but global effects through a short pathway and slower but more detailed processing through a long pathway) can also be used as an explanation for the type of affective priming results found by Murphy and Zajonc (1993). In these experiments photographs of angry and happy faces, as well as affectively neutral control stimuli served as primes, and Chinese characters (or ideographs), which were unfamiliar to the participants, served as targets that had to be evaluated. A sub-optimally presented (non-conscious, or at least less conscious than in optimal conditions) prime congruently influenced (i.e., primed) the affective evaluation of a subsequent stimulus, whereas optimal (fully conscious) presentation showed no influence, and in Experiment 1 (Murphy & Zajonc, 1993) even seemed to result in a slight (marginally significant) incongruent priming effect. Murphy and Zajonc (1993) hypothesized that the effects in the sub-optimal condition were due to a diffuse and nonspecific influence outside of awareness due to insufficient activation of higher centers. In the optimal condition, however, higher 'cognitive' processes (e.g., incorporating a context, in which an affective response to an emotional face is not relevant) corrected for the influence of the prime. They also cited the LeDoux dual-route model in support of affective-cognitive independence and primacy of affective processing. Clore and Ortony (2000), in contrast, concluded on the basis of their analysis of the role of cognition in emotion that LeDoux' research is essentially irrelevant to the findings of Murphy and Zajonc. In their view, suboptimal affective priming reflects incomplete (semantic but not episodic) parsing of the prime in the sequence of processes possibly leading up to a full emotion, whereas processing in the direct route represents rapid response preparation. The response generated in the direct route, moreover, would only be a degenerate, or at the very least nonrepresentative, instance of an emotion.

A connectionist implementation of the dual-route model simulating the affective priming results would show that these results are, at least, compatible with the model. In many cases the relation between conceptual models and experimental results remains vague and open to alternative interpretations. An implementation of the conceptual model in a computational model forces one to specify many other aspects of the model which are sometimes only of secondary relevance to the processes studied. This may lead to a further development of hypotheses and predictions that may be tested experimentally. It has, for instance, not always been easy to replicate the Murphy and Zajonc results (e.g., see Rotteveel, de Groot, Geutskens, & Phaf, 2001) and the modeling work may specify exact conditions when it can be replicated. If we, moreover, succeed in simulating these results in a relatively simple model, this forms a further argument against the claim (cf. Clore & Ortony, 2000) that fear conditioning and affective priming involve different parts of the nervous system. If evolution has provided for a fear-conditioning circuit of which we have shown that it can also perform affective priming, then it would be very unparsimonious to add a new circuit for a similar affective task. The ability to perform both types of tasks, moreover, seems to have evolved from the same selection pressures. The fast detection of threat, or the absence of threat, allows the organism to prepare itself in such a manner that with fuller processing it needs less time to act appropriately. A fear preparation to a happy stimulus, for instance, could lead to the organism missing out on advantageous situations, such as additional food or social support. These common origins and the common mechanisms, which we want to show in the present study, make it unlikely that fear conditioning and affective priming are subserved by different neural circuits.

A network model was constructed that had a similar architecture as the fear-conditioning models of Armony et al. (1995) and of den Dulk et al.(1999). Both models had a modular design with within-module inhibition and between-module excitation. A direct labeling of the modules with neuro-anatomical regions, as in the fear-conditioning models was avoided in the present model. Because this work is less neurobiologically motivated than the fear-conditioning work we cannot be completely sure which parts of the cortex are, for instance, involved in the indirect route. Also the visual nature of the stimuli, as opposed to the auditory nature in the conditioning experiments, may implicate other regions in affective priming than in fear conditioning. We would, however, preserve the general distinction between a subcortical direct pathway and a, partly cortical, indirect pathway. The correspondence between the simulation results of the two fear-conditioning models, furthermore, demonstrates that implementational details such as activation rules and learning rules do not matter much. We choose to continue with the rules and parameters of the model by den Dulk et al. (1999; see also Murre et al., 1992) In the first simulation we demonstrate how the Murphy and Zajonc (1993) effect could come about within a dual-route model. In Simulations 2 and 3 we show conditions which could interfere with this effect. In the final simulation we show that the specific weight settings in the indirect pathway needed for the correction effect could come about through learning.

The model

The paradigm with which the model was built was a competitive learning paradigm, which used the activation rule from the CALM paradigm (Murre et al., 1992). An advantage of this paradigm is the gradual change of the activation over the iterations. With a stationary input the activation of the nodes keeps changing over a number of iterations until it reaches a constant level of activation, when net input to the node and decay of activation are in balance. The gradual chance of activation allows us to represent convergence time as a measure of response latency.

The underlying theoretical assumption of the model is that sub-optimal stimuli only influence affect through the direct route, whereas, the optimal stimuli are also processed along the indirect route. The model is schematically depicted Figure 27. The rough outline of LeDoux' (1996) model, on which this model is based, can be recognized. From Input there is a direct connection to the Affect-Module, which corresponds to the direct route to the Amygdala in LeDoux' (1996) model. Enhancements and corrections of this direct affect are possible through the Indirect-Module, corresponding to LeDoux' indirect cortical route. We do not claim the processes underlying the Murphy and Zajonc (1993) effect corresponds exactly to the processes in LeDoux' (1996) experiments, but do not exclude it either. The main motivation for choosing LeDoux' architecture as a basis is that it can be applied to this new data set without too many changes. Also, we see it as an advantage that LeDoux' (1996) research has shown that such an architecture plays a role in at least some aspects of cognition, so maybe also in others, like affective priming.

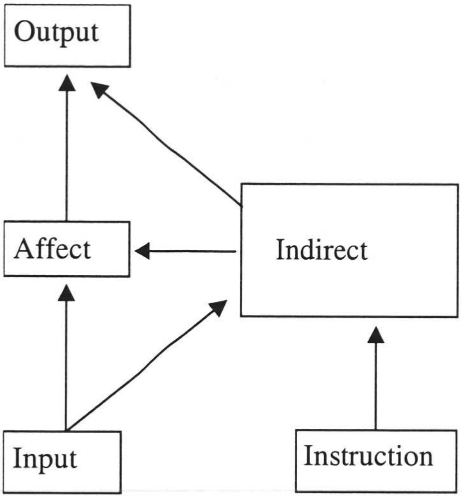

Figure 27. Global structure of the model.

A more detailed representation of the model is given in Figure 28. Stimuli are fed to the network through the pattern modules. These have no theoretical significance, but only function as a way to present patterns to the network. The Input-Group consists of two sub-modules which correspond to the two types of input presented in the experiment of Murphy and Zajonc (1993), i.e., Chinese characters and faces expressing emotions. In the model we represented only positive and negative faces for simplicity. There were also two representations for Chinese characters of which one had a slightly positive and the other a slightly negative affective value, representing an a-priori preference for some of the Chinese characters. In the actual experiment the characters were considered neutral, but in practice it is probably impossible to select perfectly neutral characters, therefore, it was considered better to average the results over slightly positive and negative characters. A further type of information presented to the network is the instruction to evaluate either the faces or the characters. The instruction is implemented as the pre-activation of a set of nodes which have representations that are compatible with the instruction. The Instruction-Module has two nodes, which stand for the indirect instruction to affectively evaluate the Chinese character, or the direct instruction to evaluate the faces. The Affect-Module, finally, consisted of two nodes, one node representing the positive affect, the other the negative.

In the modules of the indirect pathway every face and character input project towards two nodes. One node representing the high activation of the concept, which we call the congruent node, and one node representing low activation of the concept, which we call the incongruent node. By explicitly representing absence of a concept through activation of an incongruent node we can have active inhibition of other nodes that correspond to activation of the concept. The congruent face representations have positive weights to the nodes in the Affect-Module which represent the corresponding affect, and the incongruent nodes have positive weights to the opposite affect nodes, (represented by the dotted lines in Figure 28). The function of the incongruent face nodes is to correct for activations from the direct route whenever they are irrelevant. The congruent character nodes, however, have positive connections directly to the corresponding output nodes. The incongruent character nodes have no positive connections to other modules, because there is no direct affect, which needs to be corrected. Their function to inhibit the congruent character nodes through lateral inhibition in case they are irrelevant.

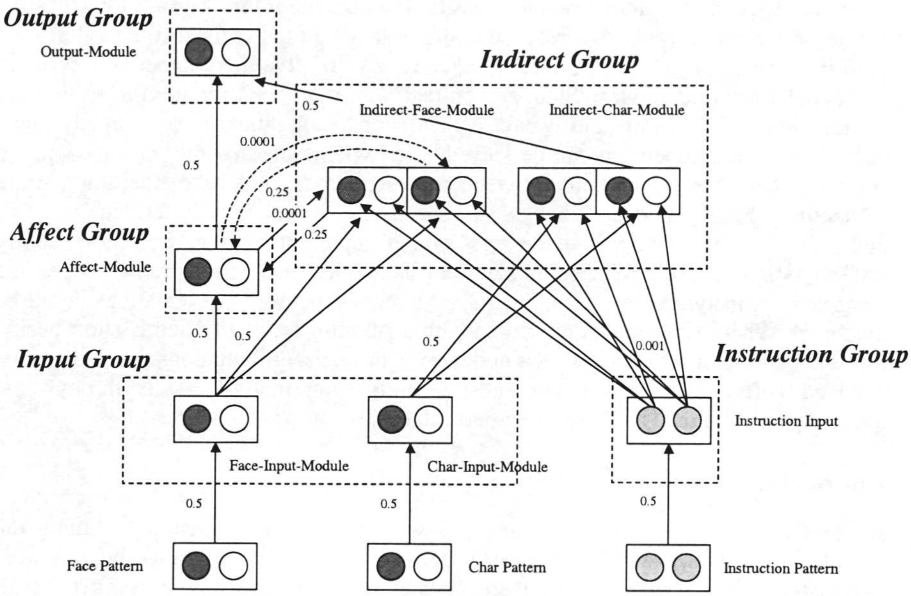

Figure 28. Boxes represent modules (i.e., groups of nodes). Arrows between circles indicate connections between nodes. Arrows between boxes indicate one-to-one connections between the nodes in the modules. The dotted arrows between modules indicates a crossed connection scheme (i.e., from node 1 to node 2 and from node 2 to node 1). All modules have inhibitory connections between the nodes within a module, these connections are -5.0 for all modules, except for the Affect- and Output-Modules, which have inhibitory connections of -0.1.

Whether the network treats an input stimulus as relevant or not depends on the instructions, and the instruction input serves to bias the activations in the Indirect-Module either in the congruent or the incongruent direction. When, for instance, the instruction is to judge the characters, activation from the Instruction-Module goes to the congruent character nodes and to the incongruent face nodes. The congruent nodes both receive activation from the input modules, but because of the strong lateral inhibition within a module, only one of them can win the competition. The small instruction activation provides sufficient bias to resolve the competition. To prevent deadlocks between two competing nodes a small noise activation was added to the incoming activation during each iteration to each node in the network. The noise was randomly chosen from a flat distribution between -0.0001 and 0.0001.

The idea of presenting stimuli suboptimally is that the very short presentation will be too short to reach higher cognitive systems and will only be processed by some lower level systems. This seems intuitively true from our own subjective experience, because it is obviously hard to detect very briefly presented stimuli, whereas they may still influence our behavior in experiments (e.g., Greenwald, Draine, & Abrams, 1996). From a modeling perspective such a mechanism is not immediately obvious. In most neural network models a brief presentation will still lead to output activation. This output will not be weaker, it will just be brief. More layers of processing could cause the activation to take more time before it reaches the output layer, but it would still reach the output layer. In the competitive modules we used in these simulations, however, it takes time to resolve the competition between two equally activated nodes. 'Sharper' contrasts are, moreover, required to solve such competition. With short and weak presentations, activations may already have decayed before competition can be solved. We hypothesize that the higher cognitive centers make finer distinctions between representations, and have, therefore, more competition among nodes. When a stimulus is presented, it will activate multiple nodes in the higher cognitive centers, which represent similar, yet more subtly different stimuli. In this specific model there is mainly such competition between the congruent en incongruent nodes, but many other competing representations could be imagined. When the model is presented with a stimulus for a sufficiently short period this will cause competition between nodes with similar representations that will not be resolved before the presentation ends. Therefore, corrective effects through the indirect pathway are less likely when presentation is short.

Connections

All of the model weights in the network were initially set by hand. In tuning the model a number of choices had to be made. We will explain here how these choices came about, and which aspects were considered. The first choice was to set all excitatory weights to the arbitrary value of 0.5. For the inhibitory connections, connections were required which were strong enough to prevent transmission of activation before the resolution of competition. Inhibitory weights of -5 proved sufficient for this. Additional tuning was necessary to allow the indirect route to correct for activation coming from the direct route. First, the inhibiting connections in the Affect- and Output-Module had to made weaker, because otherwise the inhibition of the node that was first activated would completely prevent subsequent activation of the other node. The inhibitory weights in the Affect-Module were, thus, set to the much lower value of -0.1. Second, the weights from the incongruent nodes to the Affect-Module were very sensitive; too low values were not sufficient for correction, too high values caused an extreme reversal of the effect, they were set at 0.25. Because the Chinese characters were assumed to also have some small intrinsic affective value a positive and negative Chinese character were distinguished. In the network this difference was expressed by positive connections from the character-nodes to either the positive or negative output node. These weights were set to the default value of 0.5.

There were further weights from the Instruction-Module to the Indirect-Group. The instruction input determined whether the congruent or incongruent node would win the competition. When the instruction was to name the faces, the face congruent and the character incongruent node should win. With the instruction to name the character this should be reversed. A small connection strength was assigned to it (0.001). A similar small value was given to the feedback connection from Affect-Module to Face-Indirect-Module (0.0001). The function of these connections was to make sure that the congruent node would win in case no instruction was given at all. Without these connections there would be an about equal chance of obtaining congruent or incongruent affective reactions to faces, which is now slightly biased towards congruent reactions due to the weak connections. In the experiment however there were no situations where there was no instruction so in the simulations of the experiments this weight did not contribute to the results.

Simulation 1: The Murphy and Zajonc (1993) effect

Four conditions of Experiment 1 of Murphy and Zajonc (1993) were simulated: positive and negative sub-optimal primes and positive and negative optimal primes. In a single trial first the prime was presented by clamping one of the Face-Input-Module nodes at 1.0. This presentation lasted 5 iterations in the sub-optimal condition and 20 iterations in the optimal condition. The presentation time of 5 iterations for the sub-optimal conditions was chosen so that it could activate the Affect-Module through the direct connections, but did not lead to resolution of competition in the relevant sub-modules of the Indirect-Module. Immediately after the prime, the target was presented by clamping one of the character-input-nodes to 1.0 for 20 iterations. The activation of the output node was recorded after the presentation of the target. The instruction to evaluate the character was activated during the whole trial.

The evaluation was determined by taking the difference in activation of the two output nodes (positive output activation minus negative output activation). The activation of the nodes could vary between 0.0 and 1.0, so the affective evaluation could vary between -1.0 and 1.0. Because Chinese characters also had some intrinsic affective value, in every condition both positive and negative characters were presented. For all four conditions output was averaged over 11 presentations of a slightly positive character and 11 presentations of a slightly negative character.

A regression fit (the slope was 14.6 and the intercept 3.1) was performed between the activations of this simulation and the ratings from Experiment 1 of Murphy and Zajonc (1993). This has no theoretical significance but facilitates comparison between simulation and experiment. All other simulation results in the paper were transformed according to the same regression fit.

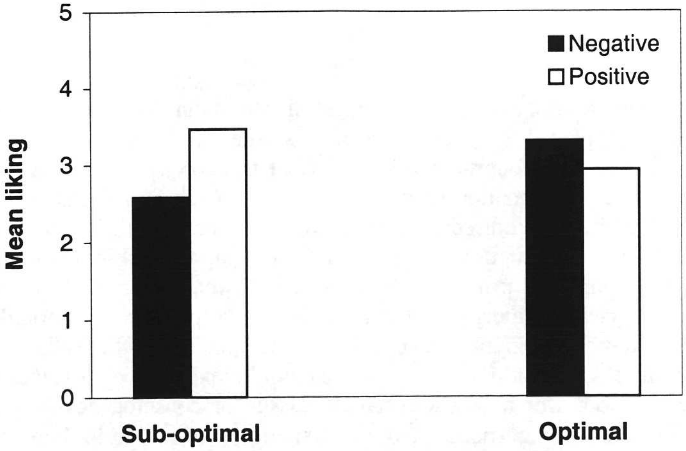

Figure 29. Ratings of the Chinese ideographs in the four priming conditions obtained from the network in Simulation 1 after a linear transformation of the activation.

Results

The judgements of the model are shown in Figure 29. A T-test showed a significant congruent difference in the sub-optimal condition ($t(10) = -94.6$, $p < 0.0001$). The incongruent difference in the sub-optimal condition was also significant ($t(10) = 46.1$, $p < 0.0001$). Overall, the simulations produce results very similar to the results of the experiments, demonstrating that the dual-route architecture is, in principle suitable for explaning the Murphy and Zajonc results. This is, of course, no more than an existence proof. It remains possible that these results can also be obtained with other types of architectures.

Simulation 2: Instruction to judge the prime

Published experimental results usually concern striking effects, which deviate from the expected. Model development in psychology is often driven by such remarkable results because they receive much attention in the literature. Only when some implicit null-model seems contradicted, it is deemed necessary to construct some explicit (i.e., computational) model to account the deviating findings. It is important, however, not only to design models that capture special aspects of behavior, but also address the more common aspects of expected behavior. In the following simulation we instructed the network to evaluate the prime and ignore the target. Such direct instructions have, for instance, been applied in an experiment by Geutskens en Rotteveel (1996). This simulation was equal to the first, but now the name-face-node of the Instruction-Module was clamped to 1.0, instead of the name-character-node. The results of this experiment are not entirely trivial because we ask the participant to judge a stimulus which is not consciously perceived. If there is nevertheless an influence of these sub-optimal primes in the indirect instructions, why should we not also expect an effect with direct instructions? If also congruent priming is found with sub-optimal presentation and direct instructions, this means that the prime may not be perceived consciously but that it is available for evaluations when a conscious reference is made to it.

Results

The affective evaluations, after the same regression fit as in Simulations 1, are shown in Figure 30. As expected the judgements are congruent in both the optimal and sub-optimal conditions. The congruent priming effect is, moreover, stronger in optimal than in sub-optimal presentation conditions and also much larger than with indirect instructions. Direct and indirect pathway work in the same direction with this instruction. Any contamination of indirect processing in sub-optimal conditions strengthens the congruent priming effect instead of weakening it. Without indirect pathway the congruent priming effect in the sub-optimal presentation should have the same size with direct and indirect instructions. This simulation, thus, also shows that, in the model at least, all indirect processing cannot be exhaustively excluded by sub-optimal presentation, which is a longstanding issue of consciousness research (see Merikle, 1992). Both the model and the experimental research, however, show qualitative differences in processing between optimal and sub-optimal processing, which cannot be exclusively identified with conscious and nonconscious processing but probably represent different mixtures of it. A further implication of these simulations is that, if the participants, or only some of them, accidentally follow a direct instruction instead of an indirect instruction, this would strongly reduce the effect. This could be a problem in replications of the Murphy and Zajonc (1993) effect (e.g., see Rotteveel et al., 2001)

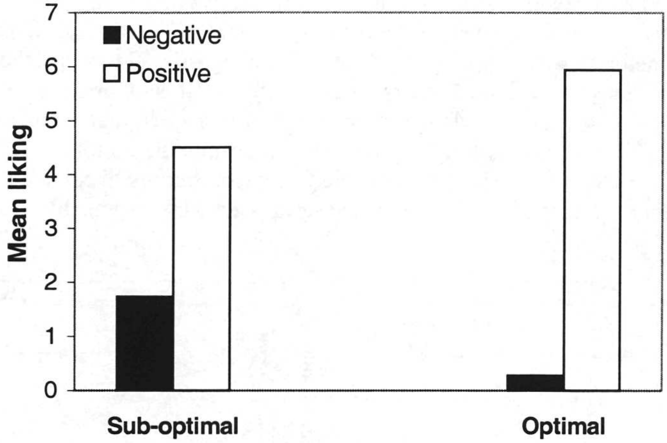

Figure 30. Direct evaluations of the prime in Simulation 2 after optimal and sub-optimal presentation.

Simulation 3: Effects of arousal in the model

It is likely that participants respond differently in the stressing context of an experiment from how they would in a more relaxed situation. Experimental stress may be particularly relevant when affective processing is investigated. How would such a state of arousal affect the outcome of an experiment? The famous Easterbrook (1959) hypothesis, for instance, argues that attention is narrowed only to the most relevant features of the environment under stressful conditions. If the prime is relevant, priming would be stronger due to experimental stress, but if it is not relevant priming may be diluted. In optimal conditions the incongruent effect may also reverse to a congruent effect due to the attentional narrowing. Experimental stress may, thus, be another factor which reduces the chances of finding the Murphy and Zajonc results. In this simulation we investigated what this model would predict if there was arousal in the participants. Our method to simulate arousal is based on the idea that increased arousal leads to the release of neuromodulators, which cause a higher level of lateral inhibition (see Izquerdo & Medina, 1997; Mintz, Gotow, Triller, & Korn, 1998, for chemical data) (see Keeler et al., 1989; Phaf et al., 2001; Rumelhart, 1997, for a connectionist perspective). In fact, raising lateral inhibition in an array of nodes also has the effect of enhancing contrast (Cornsweet, 1970) which leads to higher activation of 'central' features and lower activation of peripheral features, which may correspond to some form of the Easterbrook hypothesis (Easterbrook, 1959). The inhibition level of the Affect-Module was raised by a factor 10 in the first simulation (Simulation 3.1), and 100 and 1000 in Simulations 3.2 and 3.3. In this simulation we used the normal indirect instruction again, i.e., the name-character node was clamped.

Results

There was a congruent priming effect in both optimal and sub-optimal condition for all three inhibition levels (see Figure 31). The direct effect in the suboptimal conditions was little affected by the increase in inhibition. The indirect effect that previously led to a reversal in the optimal conditions, was no longer able to override the direct effect. The (congruent) effect in the optimal priming conditions was now also determined primarily by direct processing. Though the claim that the action of the indirect pathway is disturbed by stress seems well supported experimentally (LeDoux, 1996), to our knowledge, this is the first model that simulates such a phenomenon. The account in terms of variable within-module inhibition is, of course, rather simple and readily leads to the expected behavior, but it still cannot be excluded that some other mechanism in the nervous system is actually responsible.

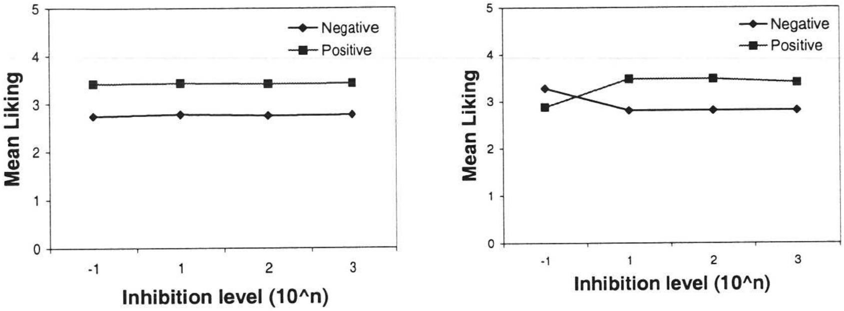

Figure 31.
a) Evaluation of the Chinese Ideographs with optimal priming in Simulation 3 as a function of inhibition in the Affect-Module. The first data-points (with inhibition -0.1) are from Simulation 1.
b) Evaluation of the Chinese Ideographs with suboptimal priming in Simulation 3 as a function of inhibition in the Affect-Module. The first data-points (with inhibition -0.1) are from Simulation 1.

These results could also give some further indication why the Murphy and Zajonc (1993) results are sometimes difficult to replicate. Only when all participants follow the indirect instructions and are relatively relaxed may the larger sub-optimal than optimal priming effect of Murphy and Zajonc (1993) be found.

Simulation 4: Learning the indirect correction

It is likely that, as (Darwin, 1965) thought, there is also a genetic basis, not only for the production, but also for the recognition of facial expressions. In a study of Serrano, Iglesias and Loeches (1992) it has been show that even very young children respond differently to angry and happy faces. Such a genetic basis, of course, does not exclude further specification and modulation of face recognition by later learning.

In terms of the dual route architecture the evolutionarily prepared reaction would, of course, be coded in the direct pathway and for the modifications (i.e., both correction and enhancement) through learning the indirect pathway should be primarily responsible. Because the optimal correction effect in the Murphy and

Zajonc results was previously attributed to the indirect pathway, it would be expected that this effect also comes about through learning.

The theoretical idea behind this model assumes that processing in the evolutionarily older direct route is primarily based on genetic factors, whereas processing in the evolutionarily newer indirect route seems primarily environmentally determined. Through the direct route it could, thus, be learned that in some contexts angry faces have a different meaning from our natural, direct, interpretation. When they are presented on a computer screen or television, for instance, they do not have the same significance as when encountered in real life. Of course, the direct route is not completely fixed; the experiments of LeDoux showed that it is capable of learning new associations. It is also known, however, that some stimuli can be more easily conditioned than others (Öhman, 1994). Context dependent conditioning, furthermore, where the predictive value of the stimulus depends on the context, also requires an indirect pathway in order to be learned. The hippocampus, which is one of the 'modules' forming part of the indirect pathway is the primary structure known to be involved in this (LeDoux, 1996).

In this simulation we will attempt to demonstrate that the weights in the network can be adjusted through a learning process so that it can correct the direct congruent priming effect. First, there will be a learning phase, and subsequently the learned network will be tested on the affective priming experiment, identical to Simulation 1. The weights from the Indirect-Module to the Affect-Module were adjusted through a learning algorithm, but all other weights remained fixed. The learning rule from the CALM module was applied in the simulations, which is a competitive learning procedure, developed by Murre et al. (1992).

$$\Delta w_{ij}(t+1) = \mu a_i \left[(K - w_{ij}(t))a_j - L w_{ij}(t) \sum_{f \neq j} w_{if}(t) a_f \right] \qquad (1.1)$$

Where $\Delta w_{ij}(t+1)$ is the weight change of the weight from node j to i. The activation is represented by a. The size of the weight change controlled by parameter μ. K and L are constants set to 1.0. The term $\sum_{f \neq j} w_{if}(t)$ is the summed activation coming from the other learning weights to node i. This term is to normalize the weights. The learning weights were all initialized to 0.0, so that before learning the network responds congruently to faces in both conditions. The learning parameter was set to 0.005.

In the learning phase we presented four different input patterns. The two types of input crossed with the two types of instruction input. The instruction input should in this experiment be interpreted as a more general context, indicating whether the face inputs are relevant or not. Supervised learning is achieved by clamping one of the affect nodes in the Affect-Module. In the relevant instruction condition the congruent affect node is clamped, and in the irrelevant instruction condition the incongruent affect node is clamped. Thus, in the irrelevant conditions the associative learning mechanism will cause the weight from the incongruent nodes to the opposite affect nodes to become stronger. Four thousand patterns were presented for 4 iterations. In each presentation one of the four patterns was chosen randomly.

Results

Figure 32 shows the weights between the Indirect- and Affect-Module. The weights roughly reach the same configuration as the hand-designed network. Weights which were not present in the hand-designed network also stay close to zero after learning in this simulation. The most crucial weights, responsible for the correction effect, reach a value close to 0.25. Overall, the results of the affective priming simulation as shown in Figure 33, have a very similar pattern as in Simulation 1. Again this simulation does not show that the correction effect in humans is caused by learning, but it strengthens the idea that it is. Though little is known of affective priming in children, it suggests that congruent priming should be found in all ages, but that an incongruent priming effect may only develop at later age. Again this suggests that the Murphy and Zajonc results may not replicate in individuals who have not learned to correct for conscious affect.

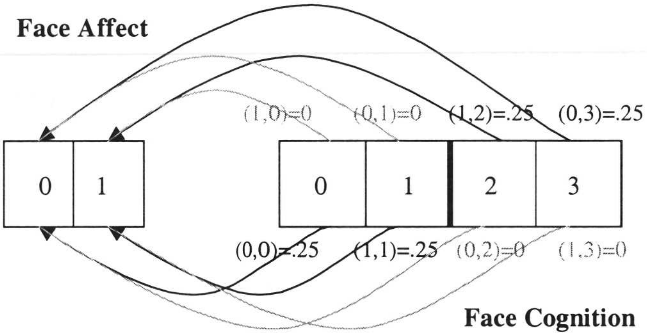

Figure 32. The individual weights between the Face-Indirect-Module and the Affect-Module after learning in Simulation 4. Next to each connection is the index of the Affect-Module and the index of the Face-Indirect-Module and the connection strength it converged to.

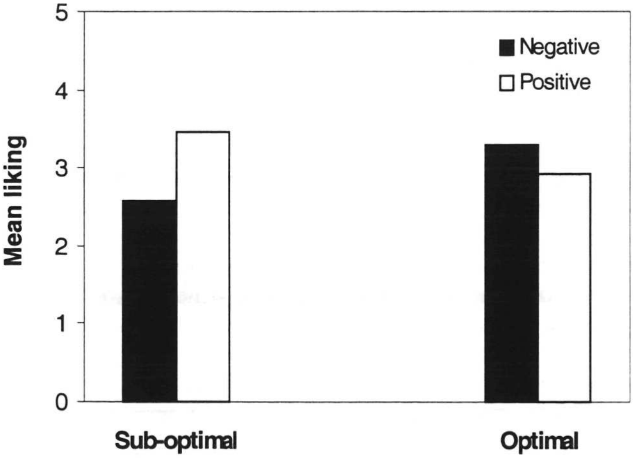

Figure 33. Affective evaluations of the network after learning.

Conclusions

Simulation 1 showed that the basic Murphy and Zajonc (1993) effect could be obtained within a dual-route architecture. Simulations 2 and 3 gave an indication of how this effect could be disturbed by high arousal and by more direct instructions. Simulation 4 showed that the weight configuration of the neural network could also come about as a result of a combination of learning and genetic predisposition. These simulations supported the idea that the underlying mechanisms of the Murpy and Zajonc (1993) effect are based on a dual route architecture similar to LeDoux' dual-route model for fear conditioning. Though such simulations in no way prove that affective priming is performed by the dual-route architecture, they make the assumption that fear-conditioning and affective priming take place in different parts of the system (Clore & Ortony, 2000) less plausible. The model also proposes an explanation for why the congruent priming effect is more robust than the incongruent priming effect. Essentially, this boils down to the fact that congruent priming is largely an innate phenomenon, whereas incongruent evaluations require elaborate learning and processing which are much more easily disturbed by a number of different factors. It should also be noted that the correction effect only forms in the presence of a biologically prepared direct reaction. This striking difference between affective and non-affective processing (i.e., in the conscious-nonconscious contrast, Murphy & Zajonc, 1993), which seems to have been ignored by Clore and Ortony (2000), can, thus, easily be accommodated in the present framework by assuming that the biologically prepared dual-route architecture is available only for affective processing.

An alternative explanation of the incongruent effect in the optimal condition could be that the participants suspected that the purpose of the experiment was to bias their judgements of the ideographs by the emotional prime, and they explicitly tried to counteract any influence of the prime. Such a more rule-based strategy is different from the learned correction effect in our model because it does not necessarily assume biologically prepared action tendencies. There are, however, findings contradicting this explanation. Rotteveel et al. (2001), for instance, found that the larger suboptimal-than-optimal pattern can also be obtained with more central measurements than a Likert scale, such as facial EMG (i.e., of the musculus zygomaticus major and musculus corrogator supercillii). If higher cognitive strategies played a role, we would not expect to also see them play a role in these facial muscles.

An even more central measure is provided by the recent neuro-imaging methods that have already been used frequently to study the neural processing of emotional faces. The neuro-imaging (i.e., PET) study that came closest to obtaining the larger suboptimal-than-optimal pattern was performed by Morris, Öhman, and Dolan (1998; 1999). The difference in activity between angry faces that had either been previously conditioned or not was larger in suboptimal than in optimal presentation conditions for the right amygdala, whereas the reverse was true for the left amygdala. Covariation techniques showed that, only with masked (suboptimal) presentation, right amygdala activity correlated positively with thalamic activity (i.e., the direct route), but correlated negatively with orbitofrontal activity (i.e., the indirect route). Also with masked presentation of fearful and happy faces, that were not conditioned beforehand, the right amygdala was clearly implicated in a fMRI study (Whalen et al., 1998). This evidence not only strongly supports, at least for the right amygdala, the LeDoux model, but also supports the idea that the primary function of the indirect pathway is to inhibit affective reactions.

The present model may be one of the very few computational models to actually simulate the fate of suboptimally presented stimuli. These stimuli are not processed completely in the indirect route because due to the short presentation time the competition cannot be solved before the activations have decayed. Small activations suffering from the inhibition by competing representations, can, however, still exert some influence over subsequent processing (e.g., priming). Processing in the direct pathway is little hampered by suboptimal presentation and this processing is primarily responsible for the congruent affective priming. Another somewhat similar account of suboptimal presentation is the object substitution theory by Enns and Di Lollo (2000). In this theory the function of lateral inhibition is taken over by re-entrant (or recurrent from a higher level to a lower level) connections that allow for a comparison of activity at two different levels. Presentation is suboptimal in this view when there is a mismatch between the re-entrant signal and the lower-level activity. A mismatch can, for instance, occur when either the lower-level activity has decayed due to the short duration of presentation or has been replaced by a mask.

It is likely that the two types of accounts and also the two types of connections (i.e., lateral inhibition and recurrent excitation, see also Phaf et al., 1990) should both be used to explain suboptimal processing. Both can serve to implement the type of route selection as a function of presentation condition that has been simulated here. The inclusion in the model of recurrent connections, only in the indirect pathway, would probably have enabled more accurate simulations at the expense, however, of making the model more complex. In a broader sense, however, the dual-route architecture itself can be envisaged as a (dual-route) implementation of object substitution. A strong congruent reaction is only observed when direct and indirect processing match. A weaker congruent reaction occurs when there is no re-entrant (i.e., indirect) signal, and, finally, an incongruent reaction (i.e., only the 'mask' is seen) is given in non-matching conditions.

Though a model like the present one can only model nonconscious processes, important aspects of the contrast between conscious and nonconscious affective processes are captured by the model. Nonconscious processes in the direct route are largely fixed due to the biological preparation and do not seem to allow for much modulation by other nonconscious processes. Nonconscious processes also occur in the longer indirect route before competition is solved (i.e., before relaxation, see Kihlstrom, 1987) and the winning activations can reach working memory (Phaf & Wolters, 1997). The latter nonconscious processes differ from the first type by their ability to support all possible parsings (e.g., both congruent and incongruent responses) of the stimuli. Conscious processes are characterized by a definite choice from all these alternatives (e.g., only congruent-negative or incongruent-positive). Modelling of the kind reported here, therefore, supports experimental efforts to contrast conscious and nonconscious conditions (Merikle, 1992) by providing a detailed analysis of the component processes, so that the properties of conscious processes can eventually be combined in a bottom-up fashion. Two types of dissociations between conscious and nonconscious processes are distinguished here. Dissociations within the indirect pathway can occur both for affective and non-affective processing. Dissociations between direct and indirect pathways may be specific for affective processing. In the question what distinguishes emotions from other mental phenomena, this study takes the position that only the core components of emotions, such as autonomic and bodily states and states of action readiness, can be (pre-)activated through direct nonconscious processes in a subcortical pathway.