The Use of an Implicit Grammar Task and Eye Measurements to study the Somatic Marker Hypothesis

FINAL REPORT NOVEMBER 2004
BIAL grant 21/02

The heart has its reasons which reason does not know
Pascal, *Pensées*, 1670

This final report covers the first pilot experiment and the second formal experiment. The results are exciting and suggest that the new methodology developed in this project is able to uncover the subtle processes involved in intuitive decisions in complex situations.

PROJECT TEAM:

Prof. Dr. Dick J. Bierman, University of Amsterdam
Dr. Axel Cleeremans, Université Libre Bruxelles
Drs. J. van Ditzhuyzen (research assistant), University of Amsterdam
Simon van Gaal (student), University of Amsterdam

CONTENTS

Abstract

This project seeks to validate Damasio's somatic marker hypothesis with an artificial grammar implicit learning task. Damasio's theory serves as a theoretical framework to study and differentiate the intuitive decision-making process. In the pilot experiment thirty-three subjects are instructed to choose one pseudoword out of two pseudowords. One is from planet 'Mars', the other from planet 'Venus'. Both words are constructed by two different sets of rules (grammar A and B). Subjects get a reward (money gain) when choosing the word from Mars, the 'correct' word and get punished (money loss) by choosing the word from Venus, the 'incorrect' word. Because the word from Venus is associated with punishment this word is gradually negative emotionally marked. Likewise the word from Mars is gradually positive emotionally marked. Pupil size variations (temporal resolution of 33 ms) and skin conductance responses were measured while making 100 decisions. After every ten trials the subject is asked on which grounds s/he makes her/his decisions. The performance of subjects started to improve without being able to specifically formulate one grammatical rule of one of both words (implicit learning hypotheses). In this pre-conceptual phase the skin conductance response was significantly larger before making incorrect decisions, than before making correct decisions (somatic marker hypothesis). Subsequently, the pupil size of subjects was significantly larger looking at incorrect words than looking at correct words before actually making a decision, but only in the pre-conceptual period (somatic marker hypothesis). On a single-trial level a larger pupil when looking at the incorrect option than when looking at the correct option predicted making the right decision in the pre-conceptual period. The latter conclusion supports strongly the interpretation that advantageous intuitive decisions are (partly) driven by somatic markers.

In the formal experiment 53 subjects did the same grammar learning task as in the pilot study with a slightly adjusted instruction. Only eye tracking data were obtained. For about half of the subjects, the knowledge elicitation procedure was more elaborate than in the pilot. Results of this study replicated the results of the pilot study. The two levels of sensitivity of measurement of explicit knowledge yielded identical results. This supports the conclusion that implicit learning resulted in increased performance before any explicit rule could be formulated. The pupil dilation was larger when looking at the incorrect alternative and mostly so in the pre-conceptual period, i.e when the subjects had no idea how they choose the correct alternatives.

Further research should focus on the individual differences in each of the 3 aspects that apparently play a role in intuitive decision making. Implicit learning, the development of a somatic marker and eventually the ‘listening’ of the decider to this subtle bodily signal. Especially this last aspect might enable training by bio-feedback methods of a crucial part in the intuitive decision process.

General Introduction

To study intuition scientifically we must first have a clear definition of this concept. The somatic marker hypothesis could serve as a theoretical framework to study the phenomenon of intuition. It is generally accepted that intuition is about decision-making and that is a two-steps process. First, knowledge is implicitly acquired and second, dependent of outcomes from the past, this knowledge is positively or negatively marked (Damasio, 1994). Intuitive decisions result from the non-conscious use of knowledge of past outcomes of decisions and the somatic, emotional marker that is associated with this knowledge (Bierman & Cleeremans, 2004). A negative somatic marker is thought to restrict the 'search space' in working memory which causes bad options not to enter consciousness. In contrast, a positive somatic marker should cause potentially better decisions to be activated automatically which leads to making the right decision.

Damasio's somatic marker hypothesis

For a long time it was thought that decision-making was a pure rational process. However, in the last decade it has begun to be more widely accepted that emotional reactions are central factors in rational human behavior (Damasio, 1994; Lang , 1995; LeDoux, 1998; Murphy & Zajonc, 1993). Damasio's somatic marker hypothesis underlines the influential role of emotions on the workings of cognition in general, and for decision-making in particular. According to this theory decisions can never be the result of cognitive processes alone because of the incapacity of rational systems to sequentially search for fast and good decisions. A pure rational solution should require too much time considering all possible options and making a list of all possible outcomes of every option. After such an exhaustive search a cost-benefit analysis is required to eventually make the right decision. All these processes take too much time and space in working memory. Time that in real life is not available. So, to be capable of making fast and correct decisions Damasio assumes that cognitive processes are being assisted or guided by more basic emotional mechanisms. Evidence in support of this idea comes from studies of normal control subjects and patients with bilateral ventromedial (VM) frontal damage during the performance on the gambling task (Bechara et al, 1996, 1997, 2000).

The Gambling task

The gambling task is designed to mimic real life decisions in the way that it factors uncertainty, reward and punishment. The task involves four decks of cards, named A, B, C and D. The goal is to maximize profit on a loan of play money. Subjects are required to make a series of 100 card selections. Cards have to be selected one at a time, from any deck, and subjects are free to switch from any deck to another, at any time, and as often as they wish. Every time a subject selects a card from deck A or B, s/he gets \$100, and every time deck C or D is selected, the subject gets \$50. However, in each of the four decks subjects encounter unpredictable money loss (punishment). The punishment is higher in the high-paying decks A and B, and lower in the low-paying decks C and D. In decks A and B the subjects encounter a total money loss of \$1250 in every ten cards. In decks C and D the subject encounters a total loss of \$250 in every ten cards. Thus, in the long term, decks A and B are disadvantageous because they cost more. On the other hand decks C and D are advantageous because the result of continuously choosing cards from these decks is an overall gain in the end.

Damasio found that normal control subjects gradually select more cards from the advantageous decks C and D and less from the disadvantageous decks A and B. In contrast, patients with bilateral lesions in the VM prefrontal cortex do not increase the number of their selections of cards from the good decks. They persist in selecting more cards from the bad (risky) decks (see figure 1).

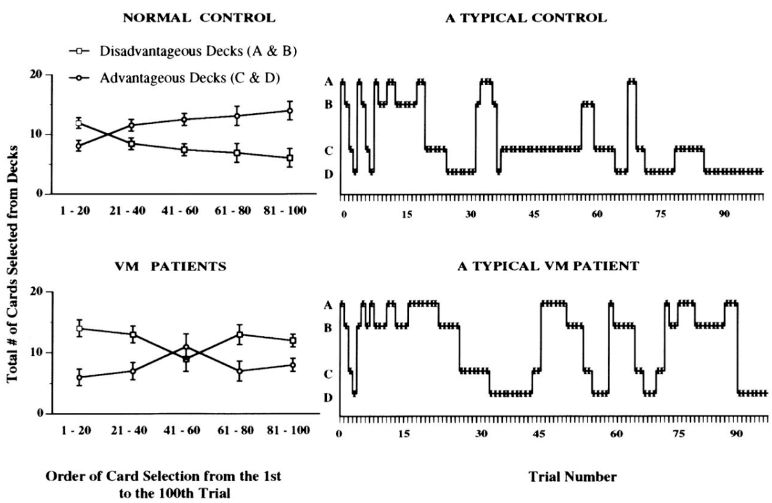

Figure 1. (Left panels) Card selection on the gambling task as a function of group (normal control, VM patients), deck type (disadvantageous versus advantageous) and trial block. (Right panels) Profiles of card selection obtained from a typical control and a typical VM patient (Bechara et al., 2000).

Normal control subjects thus learn from their previous selections and adjust their number of selections from each of the four different decks in the right manner and maximize their profit in the end. VM prefrontal patients do not learn from the consequences of decisions from the past or at least do not adjust their behavior.

The role of the ventromedial prefrontal cortex in decision-making

The prefrontal cortex is thought to be especially involved in executive processing. The ventromedial prefrontal cortex in particular seems to be an important neural structure in the decision-making process (Bechara et al., 1994, 2001, 2002; Manes et al., 2002). Different studies showed that the performance of drug addicts on the gambling task is similar to VM patients (Bechara & Damasio, 2002, Bechara et al., 2002). This could indicate that the underlying cause of addiction is a not optimal functioning VM prefrontal cortex or the other way around, drug abuse causes VM problems. Not being capable of recognizing the consequences of certain decisions could lead to addiction. In general, drug addicts are aimed at gaining immediate pleasure instead of long-term satisfaction. This is also a characteristic of VM patients. Different studies do also have shown abnormal functioning of the VM prefrontal cortex of cocaine users (Londen et al., 2000) and alcoholics (Volkow and Fowler, 2000). The question rises why VM patients and drug addicts have so much difficulty making the right decisions. Why do they not learn from mistakes from the past? Why can they not make an exact calculation of the future outcome by relating the present choice to choices and consequences experienced in the past?

Damasio's somatic marker hypothesis assumes that the VM prefrontal cortex links facts that compose a given situation and the emotion previously paired with this experience. When people face a situation for which some factual aspects have been previously categorized, the dispositions associated with this situation are activated in higher-order association areas. The result is the reconstruction of a previously learned factual-emotional set. This constrains the process of reasoning over multiple options and multiple future outcomes. The current somatic state marks the scenario as good or bad. When the outcome of this process becomes conscious, the somatic state operates as an alarm signal. The somatic state is alerting us to the goodness or badness of a certain option-outcome pair and this can lead to avoidance of the option at hand. When the outcome of this process does not become conscious, the somatic state constitutes a sort of biasing signal. The somatic state will non-consciously influence the decision-making process by inhibiting certain action tendencies. To test this hypothesis skin conductance responses were measured while doing the gambling task (Bechara et al., 1996, 1999).

Skin conductance response as somatic marker in gambling tasks

During the performance on the gambling task skin conductance responses were measured. Every time the subject picks a card, the magnitude of the SCR in the time window from 5 seconds before to 5 seconds after picking a card was measured. In this manner, three types of responses can be identified. 1) The reward SCRs, those occuring after cards with reward. 2) The punishment SCRs, those occuring after cards with punishment. 3) The anticipatory SCRs, those occuring before turning a card. In this period, the subject reflects on from which deck to choose a card. These phychophysiological experiments show that both normal control subjects as well as VM patients generate SCRs after picking a card. In this respect, both groups are thus similar. However, both groups do differ in generating anticipatory SCRs. Control subjects gradually begin to generate high-amplitude anticipatory SCRs to disadvantageous decks and not to advantageous decks. VM prefrontal patients fail to do so in respect to all four decks. In figure 2, magnitudes of anticipatory SCRs are plotted for normal control subjects and VM patients as a function of deck and card position within each deck.

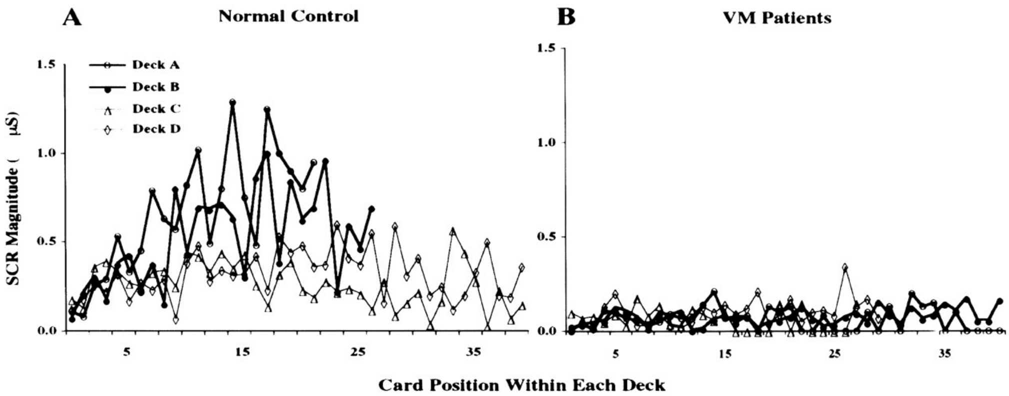

Figure 2. Magnitudes of anticipatory SCRs as a function of group (normal control subjects versus VM patients), deck and card position within each deck (Bechara et al., 2000).

What can be assumed is that a good decision generates a positive somatic marker and a bad decision a negative somatic marker. When we are facing a new decision, the earlier acquired somatic marker will be re-activated and will direct our decision to the advantageous decks C and D. An experience leading to failure or punishment generates a negative somatic marker, which will lead to avoidance of the disadvantageous decks A and B in the future.

The paradox in intuitive decision-making

If we take a closer look at the results presented earlier a paradox can be noticed. The essential measure, the anticipatory somatic marker is not in accordance with expectations stemming from the somatic marker hypothesis of intuition. The somatic marker is assumed to help people making the right decision and to avoid risky choices. Someone who 'listens' to his somatic marker should often make the right decision. However, what we have seen is that the largest somatic marker is found just before making the wrong decision (Bechara et., 1996; Bierman and Cleeremans, 2002). The most obvious conclusion would be that a larger somatic marker 'causes' someone to make the wrong decision, instead of preventing someone to make that choice. The anticipatory somatic marker is a crucial element in intuitive decision-making and these findings are difficult to reconcile with the intuitive decision-making theory. An intuitive person is seen as someone who quickly develops a somatic marker, which distinguishes between potentially good or bad decisions and also actually uses this somatic marker to make the right decision. This does not seem to happen.

If we want to investigate the cause of this paradox we should be able differentiate between somatic reactions to potential options before actually making a choice. Damasio's experiments are unable to distinguish between reflection on positively marked options and negatively marked options. In this experiment we do by using a different and faster dependent measure.

Pupil size and SCR as somatic markers in an artificial grammar implicit learning task

Emotions are accompanied by an emergency reaction, or 'fight or flight response' generated by the autonomic nervous system (ANS). Characteristic bodily signs of emotional arousal are increasing heart rate and sweaty palms. Another fairly simple indicator of ANS activity is pupil size variation. Although there are difficulties interpreting pupil size data there is one big advantage. By using pupil size as somatic marker it is possible to differentiate between somatic reactions on each of both potential options before making a decision. Although investigations of pupil size reactions to various stimuli are few, previous studies assume that pupil size is both related to cognitive processing as well as affective processing. For example Beatty (1982) has shown that pupil dilation is positively associated with increased cognitive load. The pupil dilated more under conditions of increased attention, increased demands on the working memory and while interpreting difficult material, like emotional stimuli (see also Siegle et al. 2001). However Hess (1972) suggested that there is a continuum ranging from extreme dilation to pleasing stimuli to extreme constriction to unpleasant stimuli. In contrast, Janisse (1974) suggested that pupil size is linearly related to the intensity dimension of stimuli. In this way pupil size is the largest at the negative and positive ends of the continuum of valence and smallest in the center, which represents neutral affect. Recently, Partala and Surakka (2003) have shown that the pupil is significantly larger in response to both negative and positive emotional stimuli as compared to the dilation to emotional neutral stimuli. Not all reported data are consistent and therefore more research is needed to assess this relation between pupil dilation and emotion.

This experiment is based on a previous study of Bierman and Cleeremans (2002, in press). The task used in this experiment is an artificial grammar implicit learning task (Cleeremans et al., 1998; Reber, 1967). Subjects have to choose the correct word out of two 'words'. Both 'words' are series of six symbols and are constructed by two different sets of rules (grammar A and B). The symbols], #, * and + are used (see figure 3). One word is said to be from planet 'Mars', the other from planet 'Venus'. One word of Mars and one of Venus are shown on a computer screen, right versus left. Subjects get a reward (money gain) when choosing the word from Mars, the correct word and get punished (money loss) by choosing the word from Venus, the incorrect word. Because the word from Venus is associated with punishment this word is negative emotionally marked. Likewise the word from Mars is positive emotionally marked. Correct words can only be identified by implicitly or explicitly learning the underlying grammar rules through feedback from outcomes of previous decisions. Pupil size variations and skin conductance responses were measured while making 100 decisions. After every 10 trials the subject is asked on which grounds s/he makes his/her decisions.

By using skin conductance response we try to replicate Damasio's findings (1994, 1996, 1997, 2000) with a different, more complex task. In a previous study of Bierman and Cleeremans (2004, in press) is shown that the anticipatory somatic marker is larger in the pre-conceptual (implicit) phase, than in the conceptual phase. In this study a new variation on Damasio's paradigm is introduced: the use of pupil size variation as anticipatory somatic marker. The latency of pupil size variation after stimulus onset is about 300-400 ms (Beatty, 1982; Siegle et al., 2001; Partala and Surakka, 2003), which in contrast to the latency of skin conductance (about 2000 ms) is very short. Because one word is presented on the right and the other on the left of the screen, it is possible to differentiate between anticipatory somatic markers on both stimuli by taking horizontal eye-movement and fixation position into account. By linking both anticipatory somatic markers to final choices it is possible to investigate if people do actually use their somatic markers to make the right decision.

The two experiments

This project was split in two parts. In the first part we explored the use of pupil dilation as a somatic marker with a limited number of subjects. In this pilot experiment we used also the GSR as dependent variable. We deemed it not advisable to change two aspects of the original gambling task, the task itself and the dependent variable, simultaneously. Using both measures prevented potentially negative results with pupil dilation to be un-interpretable. In the second confirmatory experiment the use of GSR was considered to be superfluous and thus we were able to test a larger number of subjects.

Both experiment used the artificial grammar learning task as the test-bed to investigate the somatic marker and the processes underlying intuitive decisions.

In the pilot experiment we simply asked subjects how they did perform the task in order to measure their explicit knowledge. Because of suspicions that this simple question was not sensitive enough we introduced in the confirmatory experiment a more elaborate questioning.

PILOT EXPERIMENT

Hypotheses

In this pilot experiment we expect subjects to

1) Implicitly learn the underlying grammatical rules and make the right decisions with above chance rate well before they can explicitly formulate one or more specific rules of one of both grammar sets (implicit learning hypothesis).

2) Show a larger skin conductance response before making an incorrect decision than before making correct decisions in the pre-conceptual period (somatic marker hypothesis).

3) Have larger pupils looking at 'incorrect' and thus negative emotionally marked words, than looking at 'correct' and thus positive emotionally marked words in the pre-conceptual period (somatic marker hypothesis).

4) The relation between final decisions and preceding somatic markers (arousal values) will also be explored.

Method and materials

Participants

Thirty-three subjects, 11 male and 22 female participated in this study. Their age ranged from 18 to 23 (mean=19.1, sd=1.67). The subjects were either friends of the experimenter, or freshmen psychology students who participated for course credits. Data from one subject was discarded before any analysis due to inattention during the session and failure to do the task appropriately.

Materials

The stimuli consisted of pairs of 'words', series of 6 symbols constructed according to two sets of rules, grammar A and B. The symbols [, #, * and + were used. For each stimulus exposure the two words were displayed on the screen. The location of the two words, one constructed according to grammar A, the second according to grammar B was randomized (left and right). The four possible symbols for the two grammars were identical, only the transition probabilities were reversed while the magnitude was always 0.25 (see fig. 3). The first symbol was also always selected at random. The self-transition probability reduced to zero after one self-transition occurred. So words having three consecutive '#' and ']' (grammar A) or '*' and '+' (grammar B) characters could not occur.

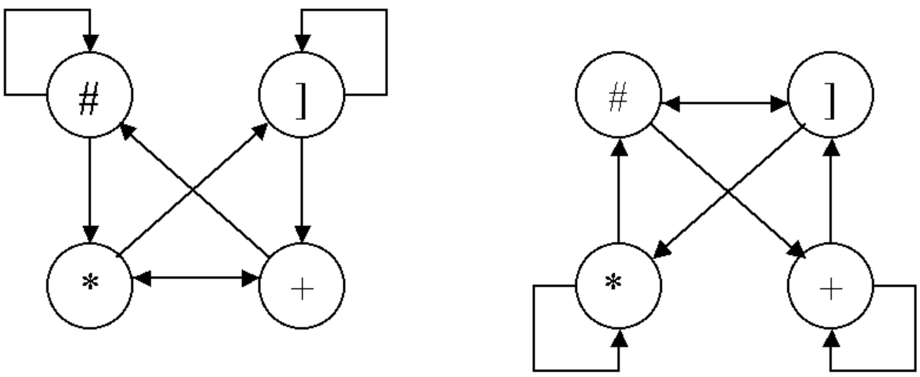

Figure 3. Transition probabilities for grammar A and B. The transition of a symbol to itself was only allowed once.

Procedure

First, the laboratory was introduced to the subject and then s/he received a written instruction form describing the goal of the experiment as a learning task and providing information about the possibility to earn prizes (see appendix 2). After this, the subject was comfortably seated in an adjustable chair in front of the computer screen. After attaching the electrodes to the middle and index finger of the left hand, this hand was positioned on the left armchair. Isotonic paste was used. The response of skin conductance on a deep breath was measured. The subject was instructed that her/his eye-movements would be measured. An adjustable headrest was used to keep the eyes at a distance 60 cm from the center of the computer screen. The eye-tracker was manually installed to exactly fit the pupil. The subject was instructed to look at a fixation cross at the center of the screen as a part of the calibration procedure. Subsequently, a demo trial was started to familiarize the subject with the type of words used in this experiment. If no questions remained, the experimenter started the experiment by clicking the mouse. This resulted in a display of the two ‘words’. The location of the correct word (left or right) was truly random. The subject could take as much time as s/he needed to determine which of the two words was from planet Mars. The experimenter stayed in the room without having a view on the display. The lighting in the laboratory was kept at a constant level for all subjects.

Feedback

After the choice was entered by a single key press, the computer marked the chosen word in black. Three seconds later the computer showed if the choice was correct or incorrect and generated a visual and auditory feedback by specifying the amount of pseudo Euros that were won or lost.

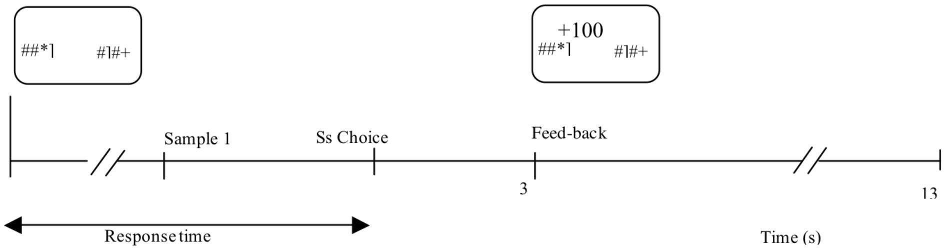

Figure 4. Timing of a single trial. Data are stored from 4 seconds before, till 13 seconds after the choice has been made.

For incorrect choices there was a 50/50 random‘punishment’ of either –10 or –100 Euros. For correct choices there was a 50/50 random reward of +10 or +100 Euros. The display showed both the reward or punishment as well as the cumulative score. Feedback remained on the screen for 6 seconds. Four seconds later the next pair of words was shown (see Fig. 4).

Elicitation of explicit knowledge

After each 10 trials, the computer generated the question: *How do you come to a choice between the two words?* The response was written by the experimenter on a standardized score form. Knowledge of the grammar was scored to have become *explicit* when the subject correctly formulated at least one rule and did not relapse to a state where the rule disappeared. Conservatively, another 5 trials were subtracted because the knowledge formulated at for instance trial 50 could have become explicit anywhere between 41 and 50, so 45 was taken as the best estimate.

Equipment

Skin conductance measurement

Two Ag-AgCl electrodes were attached to the middle and index finger of the left hand. Isotonic paste was used. The skin conductance was measured using the Orion 4AD22 which measures skin conductance using a constant AC current method (10 microamps, 100 Hz). Epochs were stored from 4 seconds before the choice till 13 seconds after the choice has been made (see fig 4). The data were sampled on interupt basis with a sample frequency of 5 samples/s.

Pupil size and horizontal eye-movement measurement

Pupil size and horizontal eye-movements were monitored with a *Viewpoint* Eye-tracker. Viewpoint Eye-tracker software is running under Windows on a PC computer. Eye-tracker data were sampled on interupt basis with a sample frequency of 30 samples/s.

Results

The implicit learning hypothesis

For each subject the start of the conceptual phase (explicit knowledge phase) was determined using the method described earlier. This was compared with their performance curve. For most subjects the performance started to increase far before they entered the conceptual phase.

Seven of the remaining 32 subjects reported an explicit rule before trial 10 (21.9%). These subjects were eliminated for further analysis because their pre-conceptual period was too short. From the remaining 25 subjects eight subjects hadn't formulated any rule after 100 trials (32%). The average trial an explicit rule could be formulated by the remaining 17 subjects was trial 58 (sd=23). The average performance curve of the 25 subjects who did have a pre-conceptual period is given in figure 5.

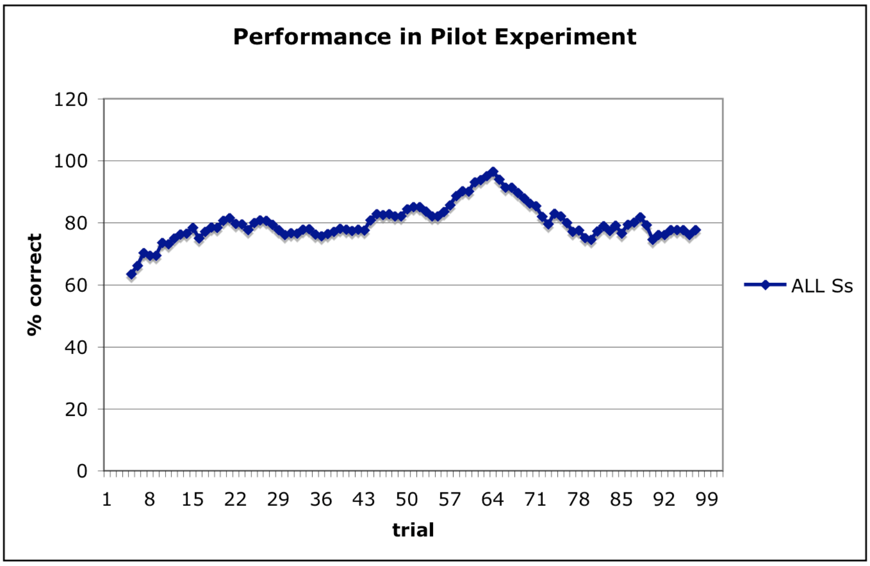

Figure 5. The moving average (9 trials) performance of the 25 subjects who formulated an explicit rule after trial 10 or couldn't formulate any rule at all.

The somatic marker hypothesis

Skin conductance responses

Due to malfunctioning of the equipment, 4 subjects had to be removed. Three of those belonged to the group of subjects who mentioned a correct rule at trial 10. The other one did never discover a rule. Baseline of the skin conductance was set to the first sample taken (4 seconds before the choice of the subject). The baseline corrected skin conductance samples were averaged over the period before feedback was given for each trial, resulting in a variable correlating with ‘arousal’ before feedback, i.e. during the decision and anticipation phase. This variable represents Damasio’s somatic marker (SM) for each choice. Subsequently, these ‘arousal’ values were separately averaged for the correct and incorrect choices per subject. To investigate the somatic marker effect in the pre-conceptual period only trials were used for which it was assessed that the subject had no explicit knowledge with regard to the grammar rules. This was compared to the results of all subjects over all trials. This resulted in the dependent variables *SM_correct choice* and *SM_incorrect choice*. Damasio's somatic marker effect was acquired by subtracting *SM-correct choice* from *SM-incorrect choice* for each subject. Both a one-sample t-test as well as a Wilcoxon signed ranks test were used for the analysis of these subtractions.

Figure 6 shows the time course of the average skin conductance in the pre-conceptual period of the 24 subjects who had a pre-conceptual period and produced valid Skin Conductance data. Figure 7 shows the time course of the average skin conductance in the conceptual period of 21 subjects who had a conceptual period.

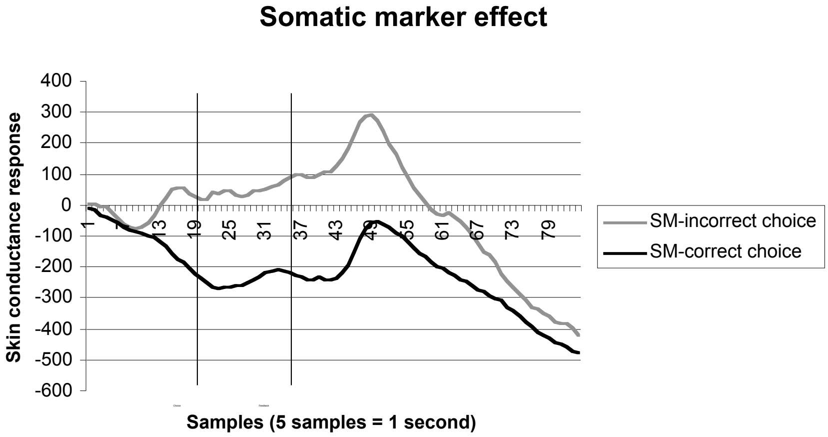

Fig 6. The skin conductance preceding, during and after feedback of correct and incorrect decisions in the *pre-conceptual period* for 24 subjects.

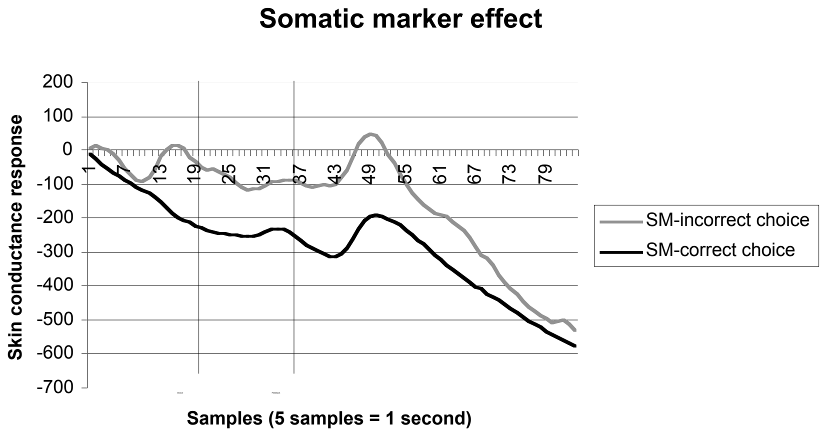

Fig 7. The skin conductance preceding, during and after feedback of correct and incorrect decisions in the *conceptual period* for 21 subjects

T-tests

The skin conductance response before making an incorrect decision was significantly larger than the skin conductance before making a correct decision in the *pre-conceptual period*, (t = 2,31, df = 23, p = 0.015 (1-tailed)). However, the *SM_incorrect choice* was not larger than the *SM_correct choice* in the *conceptual period* (t = 0.925, df = 20, p = 0.183). Over all 100 trials of all subjects the *SM_incorrect choice* was also not larger than the *SM_correct choice, (t* = 1,58, df 28, p = 0.062).

Wilcoxon signed ranks test

The skin conductance response before making an incorrect decision was significantly larger than the skin conductance response before making a correct decision in the *pre-conceptual period*, Z = 2.312, p = 0.01. This was not the case in the *conceptual period*, Z = 0,226, p = 0.821. In contrast, over all 100 trials of all subjects the *SM_incorrect choice* was larger than the *SM_correct choice*, Z = 1.697, p = 0.045.

Pupil size

By taking horizontal eye-movements into account pupil size data were categorized as *SM-correct option*, the somatic marker when looking at correct options, and *SM-incorrect option*, the somatic marker when looking at incorrect options before actually making a decision. In a pilot study the largest difference in pupil size between negative emotionally and positive emotionally marked words was seen one second after stimulus presentation. So, for both data sets data were averaged over a one second time interval (30 samples) after each stimulus onset. For a smoothed visual timeline blinks and artefacts were removed by deleting samples with a pupil aspect ratio of 0.80 or lower. Trial averages were deleted when resulting from less than 20 samples.

Figure 8 is an illustration of the pupil size of a subject who first looks at the incorrect option (about 38 samples), then looks at the correct option (about 40 samples) and finally chooses one of both words. Figure 9 illustrates which samples are used for the dependent variable in the pilot experiment. Only the first 30 samples for both options are used in the analysis. Paired sample t-tests were used for all analyses.

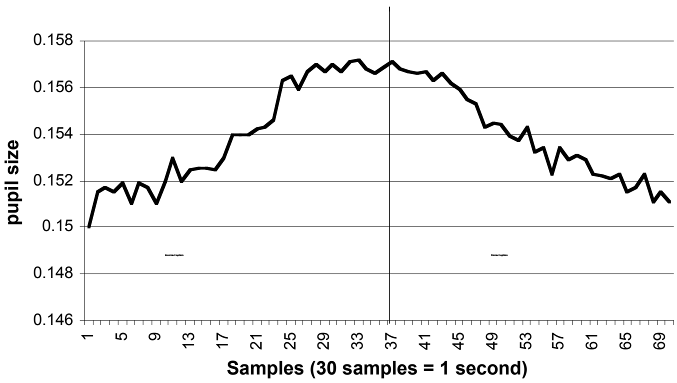

Fig 8. Example of Pupil Size as a function of time (see text)

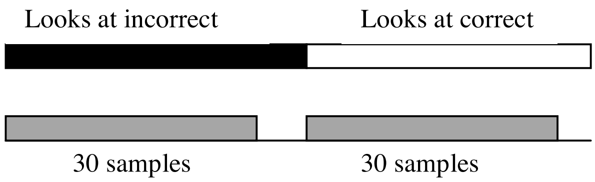

Figure 9. Although the subject looks longer to the incorrect and correct alternative only the first second of both is used to calculate mean pupil size..

The pupil size of subjects was significantly larger looking at incorrect options before actually making a decision than looking at correct options before actually making a decision in the pre-conceptual period, t = 1,87, df = 24, p = 0.037 (1 tailed).

In contrast, when taking only conceptual trials this effect disappeared. *SM-incorrect option* was not larger than *SM-correct option*, t = 1,17, df = 23, p = 0.13 (1 tailed). The mean effect size in the pre-conceptual period was 0.22 %. Means and standard deviations are shown in table 1 (pre-conceptual period) and table 2 (conceptual period).

Period	**N**	**correct**		**incorrect**		**effectsz**	**t**	**p**
		mean	**sd**	**mean**	**sd**			
Preconc	25	0.1598	0.0192	0.1602	0.0191	0.22%	1.87	*0.037*
Conc	21	0.1648	0.0160	0.1651	0.0167	0.18%	1.17	n.s.

Table 1. Mean pupil sizes split for looking at the correct and incorrect alternative for the preconceptual and conceptual period (note size is in arbitrary units).

Pupil Size and Correctness of decision

Eventually, we would like to investigate if the somatic marker actually influences the final decision on a single trial level.

If this task was no more than a simple guessing task and pupil responses should not influence the decision-making process the sum of both values of one diagonal should be the same as for the other diagonal (A + D should be equal to C + B, See below).

	Response	
Largest pupil	Incorrect	Correct
SM-correct option	**A**	**B**
SM-incorrect option	**C**	**D**

Table 2. Explanation of the formula (A + D)/ (A+B+C+D).

We can calculate a listening index as (A+D)/ (A+B+C+D) which has an expectation value of 50%. A larger value suggest that the subjects will go for the alternative that is associated with the smaller pupil size. Results were calculated for every subject for both pre-conceptual trials as well as conceptual trials. The 'listening index' for the preconceptual period was 53.5% (t = 3,02, df = 24, $p < 0.01$) but declined to 52.1% in the conceptual period (t = 0.91, df = 23, n.s.). The tables 3 and 4 below summarize the responses of all subjects as a function of the largest pupil diameter on a single trial averaged over all pre-conceptual trials and conceptual trials.

Pre-conceptual period		
	Response	
Largest pupil	Incorrect	Correct
SM-correct option	12,1 %	35,3 %
SM-incorrect option	11,2 %	41,4 %

Table 3. Final decisions of subjects as a function of the anticipatory somatic marker in the *pre-conceptual period*

Conceptual period		
	Response	
Largest pupil	Incorrect	Correct
SM-correct option	4,2 %	44,6 %
SM-incorrect option	3,3 %	47,9 %

Table 4. Final decisions of subjects as a function of the anticipatory somatic marker in the *conceptual period.*

In the pre-conceptual period, results of 18 of the 25 subjects were positive. Results differed significantly from 0 (t = 3,02, df = 24, $p < 0.01$). In the conceptual period, the results of 11 of 24 subjects were positive, 12 were negative and one was zero. These results did not differ significantly from 0 (t = 0.91, df = 23, $p = 0.37$).

FORMAL EXPERIMENT

Hypotheses

In the formal experiment the hypothesis were based upon the findings in the pilot experiment. However the skin conductance was not measured in this experiment so we expected subjects to

1) Implicitly learn the underlying grammatical rules and make the right decisions with above chance rate well before they can explicitly formulate one or more specific rules of one of both grammar sets (implicit learning hypothesis).

2) Have larger pupil size responses looking at 'incorrect' and thus negative emotionally marked words, than looking at 'correct' and thus positive emotionally marked words in the pre-conceptual period (somatic marker hypothesis). In the formal experiment all data processing was automatized and therefore the requirement for a valid measurement that the subject had to fixate longer than 2/3 secs was dropped.

3) Select more often the alternative that results in the smaller pupil response during the time preceding the decision. This might be interpreted as a direct impact of the somatic marker on the decision process.

Method and materials

Participants

Sixty-nine subjects entered in the study. Four subjects produced noisy or invalid eye tracker data. Eleven subjects mentioned a correct rule the first time they were interrogated. One subject did not understand the task. From the remaining 53 subjects, 12 were male and 41 female. Their age ranged from 18 to 51 (mean=22.7, sd=6.56). The subjects were freshmen psychology students who participated for course credits. The data of these 53 subjects were used in the final analyses.

Materials

The artificial grammar task used in the confirmatory experiment was identical to the task presented in the pilot experiment. However the equipment was significantly updated (see section on equipment). An exit interview was taken to measure if the subjects had seriously participated in the experiment (see appendix 4).

Procedure and Feedback

The procedure and the feedback to the subject of the correctness of their choice was very similar to the procedure and feedback in the pilot experiment. In the formal experiment subjects filled in an informed consent (see appendix 3). Also the instruction was slightly changed. Rather than talking about Mars and Venus, the two planets where, according to the cover story, the words came from, were now labeled X and Y. This was done because from informal observation in the pilot experiment it was found that sometimes the subjects used a judgement of the male aspect of the words (rectangular characters) to assign Mars to them. Rather than having a separate keyboard on their lap, the subjects now could respond through a separate button box. The camera of the eye tracker was placed in such a way that the subjects were less closed in. I.e. a bit more to the right side of the subject. The experiment started always with a flickering (dark-white) screen for later control for pupil size sensitivity.

Elicitation of explicit knowledge

As in the pilot experiment, after each 10 trials, the computer generated the question: *How do you come to a choice between the two words?* The response was written by the experimenter on a standardized score form. For about half of the subjects the questioning was more elaborate at trial 20 (see appendix 1 for the two methods of enquiry).

Equipment

Pupil size and horizontal eye-movement measurement

Pupil size and horizontal eye-movements were monitored with the same Viewpoint Eye-tracker as in the Pilot experiment. Data were stored with a sample frequency of 30 samples/s. The video monitor from the Pilot experiment was replaced by a LCD monitor. The keyboard which was used in the Pilot experiment to register the subjects choice of word, was replaced by a two button box which fed the responses into the parallel port of the PC.

Results

The implicit learning hypothesis

For each subject the start of the conceptual phase (explicit knowledge phase) was determined using the method described earlier. This was compared with their performance curve. For most subjects the performance started to increase far before they entered the conceptual phase.

As mentioned before, the number of subjects reporting a rule at trial 10 was 11 out of 64 (17%). These subjects were eliminated for further analysis because their pre-conceptual period was too short. From the remaining 53 subjects 22 subjects hadn't formulated any rule after 100 trials. The average trial an explicit rule could be formulated by the remaining 31 subjects was trial 49.5 (sd=26.9). The average cumulative performance curve of the 53 subjects who did have a pre-conceptual period is given in fig. 10. For each correct choice the score is 1 and for each incorrect choice the score is -1. Mean chance expectation for means score as well as for cumulative score is therefore 0.

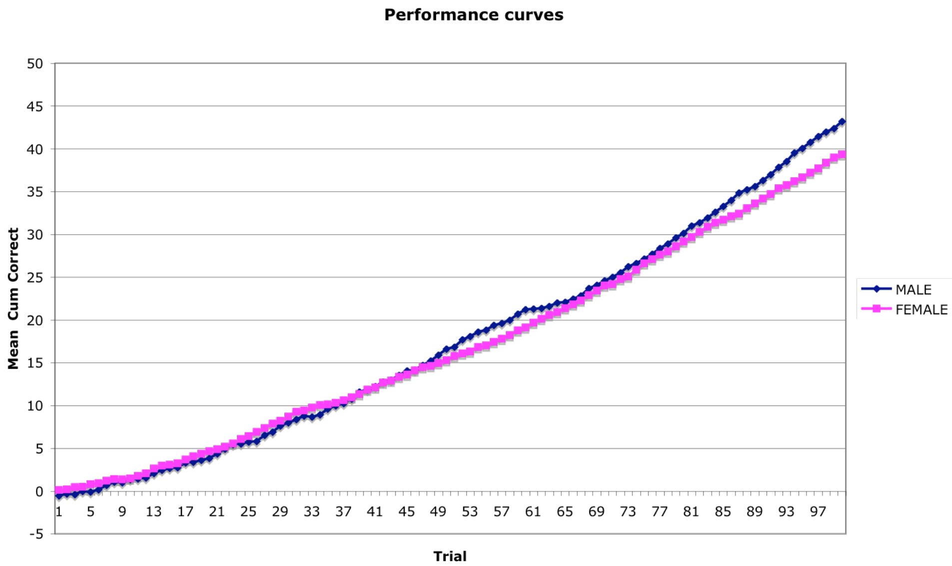

Figure 10. The cumulative performance of the 12 male and 41 female subjects who formulated an explicit rule after trial 10 or couldn't formulate any rule at all.

The performance is increasing steadily and seems to show some spurts. A moving average of 9 consecutive trials shows that there is indeed some structure in this learning (fig. 11).

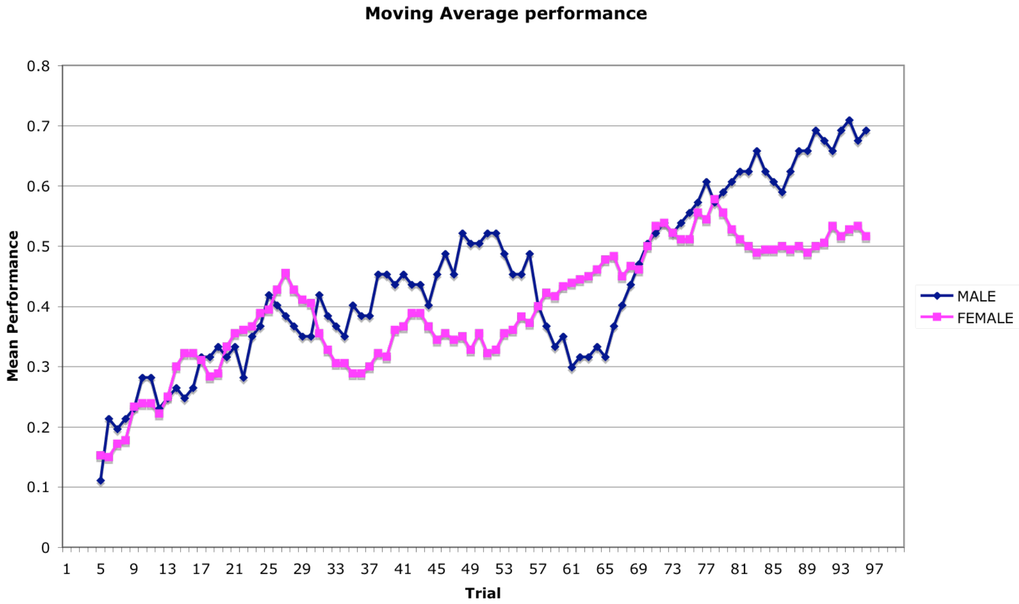

Fig 11. Moving average over 9 points of the performance of male and female subjects.

Especially the male performance has a rather sharp decline after trial 50. The female performance has a similar decline, but not as outspoken, after trial 25.

For each subject that mentioned a correct rule (after trial 10) we calculated the mean performance during the trials from 20 to 10 trials before this correct rule was mentioned the first time, i.e. in the preconceptual phase. The resulting mean score during this period of 0.312 is significantly above the expected mean chance score of 0 ($t=5.73$, $df=31$, $p < 0.0001$)

The procedure to increase the sensitivity of the measurement of when the conceptual, explicit phase started was applied to 21 of the 53 subjects. None of them formulated a rule during this interrogation. Eight even couldn't formulate any rule till the end of the experiment (38% vs 39.6% for the whole group). The average trial at which these interrogated subjects could formulate an explicit rule was 57.5, a little bit worse than the 49.5 for the whole group of subjects who formulated a rule during the experiment. There were two subjects of 21 (9.5%) formulating the first rule at trial 30, the first opportunity after the interrogation. However for the non-interrogated group this number was 4 (12.5%). The conclusion is that the more sensitive measurement did not yield any extra knowledge and moreover did apparently not result in a better discovery of rules after the interrogation.

The somatic marker hypothesis

Pupil size

Flickering of the screen resulted for all subjects in systematic variation of the pupil size. Mean pupil size for dark screens was about 0.142 and for white screen 0.132 (a.u.) The difference was of course highly significant (difference is 0.010, t= 8.23, df=51, p << 0.00001)

Eye tracker data were further analyzed in two steps. First the software categorized the direction of observation in either left word or right word. Then the mean pupil dilation for either left or right word was calculated. Because of the placement of the camera slightly at the right side of the subject's head a small but systematic error was introduced since the pupil size is projected onto the plane perpendicular to the line from nose to camera. Also a very small but systematic effect was caused by apparently slightly different lightening of the wall behind the LCD monitor. These small artifacts become apparent because independent of conditions the pupil size for words presented on the right side is slightly. We corrected the pupil sizes measured while looking at the left word by multiplying these values by the following factor

Correction factor = mean pupil size right word/ mean pupil size left word

After correction the mean pupil sizes looking at right and left words are equal as they should be.

By combining the pupil data with the stimulus presentation orders used we are able to calculate *SM-correct values*, the somatic when looking at the correct word, and *SM-incorrect values*, the somatic marker when looking at the incorrect word **before** actually making a decision

So, for both data sets data were averaged over a one second time interval (30 samples) after each stimulus onset. For a smoothed visual timeline blinks and artefacts were removed by deleting samples with a pupil aspect ratio of 0.80 or lower.

Figure 12 shows a graph of a subject's first trial data of the raw X-gaze position and the software categorization in left, right or undecided together with the pupil size data. It can be seen that the categorization starts with a value of 0.5 which means undecided (the subject is looking somewhere in the middle, very probably to his or her score which is displayed in the middle of the screen).

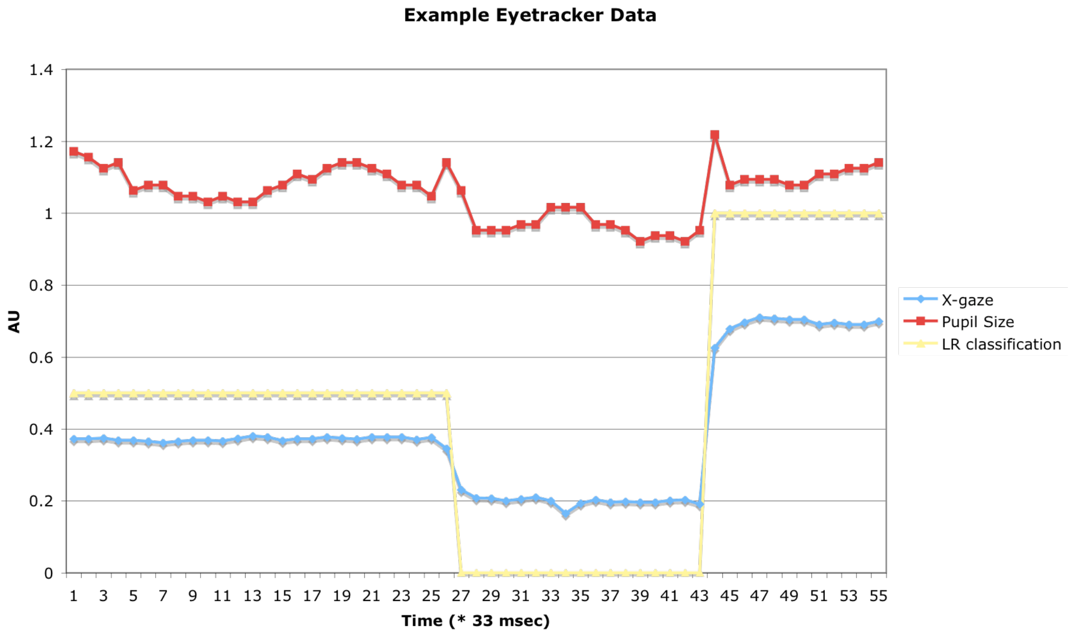

Fig 12. An example of Eyetrackerdata with pupil size, X-gaze and Left-Right classification output.

Then the subject starts looking at the left alternative for about half a second and subsequently to the right for about 400 msec before deciding to enter a response. During an eye-movement the pupil size measurement is not valid (the pupil image is blurred and a too large value is calculated). Also it can be seen that after movement from left to right the measured pupil size is larger due to the placement of the camera slightly to the right of the subject.

In table 5 the results of the mean (corrected) pupil size and the effect size as percentual difference are given.

Period	**N**	**correct**		**incorrect**		**effectsz**	**t**	**p**
		mean	**sd**	**mean**	**sd**			
Whole	52	0.1440	0.0204	0.1444	0.0204	0.27%	1.7	*0.048*
Preconc	29	0.1447	0.0204	0.1455	0.0206	0.55%	1.9	*0.033*
Conc	29	0.1450	0.0193	0.1454	0.0203	0.27%	0.7	n.s.

Table 5. Results of pupil size measurement when subjects look at the correct and the incorrect alternative preceding the actual decision

As in the pilot study the (corrected) pupil size of subjects was significantly larger looking at incorrect options (mean=0.1444, sd= 0.0204) than when the subject was looking at correct options (mean=0.1440, sd=0.0204). Whereas the effect size in the pilot study was 0.22%, the effect size in the formal experiment is slightly larger, 0.27%, but still small and still robust. (t=1.7, df=51, p=0.048 one tailed). For subjects for which a split in preconceptual and conceptual period could be made, the effect is most pronounced in the preconceptual and less so in the conceptual period. Note that the effect size in the subjects that never formulated a rule is low so that after removal of these subjects the effect size in the remaining subjects gets higher on the average.

Pupil Size and Correctness of decision

The pupil size when looking to (and as we assume reflecting on) an alternative before correct decisions shows a larger difference of 0.28% between the two (correct and incorrect) alternatives (t= 1.675, df=51, p = 0.05) than when we look at what happened before what eventually became an incorrect decision (0.21%, ns). These data suggest that when there is a more distinct difference in somatic marker during the reflection preceding the decision, subjects will make better decisions. However in order to make a more definite statement the data should be analyzed on a trial by trial basis like in the Pilot study.

Although there is ample evidence that the somatic marker as operationalized by the pupil size is correlated with the correctness of the alternative that is contemplated and there is suggestive evidence that this has an impact of the eventual decision we want to investigate more in detail if the somatic marker actually influences the final decision on a single trial level.

For each trial therefore we determined which of the two alternatives was associated with the largest pupil size (after correction for placement of the camera). Then we checked if the subject's final choice was this largest pupil size 'inducing' alternative or the other alternative. The number of times that the subject eventually choose for the other alternative was calculated as a percentage of the total number of trials. This number gives the tendency to 'listen' to the somatic marker (as a warning) in the reflection period. The mean listening index was 50.4% (n.s.). If we split these results for the preconceptual period and the conceptual period, like we did for the pilot study, we do not observe the same decline as in the pilot study. Rather, the tendency to 'listen' to the somatic marker is quite similar in the preconceptual period and conceptual period. However there was a quite interesting difference between male and female. For the 11 male subjects the mean listening index was 47.5% (t = -2.645, df=10; p=0.025). They choose significantly more often the alternative that gave rise to the largest pupil size. This result is much stronger in the preconceptual than in the conceptual period (46.4% versus 48.2%). For the female the reverse effect occurred. They choose in the conceptual period with 51.5% more often the alternative that is associated with smallest pupil size. This declined to 51.2% in the conceptual period. The results for male seem to run counter to the results obtained in the pilot experiment but they indicate nonetheless that in some way or another the subjects' pupil size when looking at the two alternatives is a predictor for the final choice that is going to be made. Further analyses could shed light on the processes that are either correlated or even causal for this connection.

Discussion of all results

From the experiments reported here, it can be concluded that subjects make correct decisions far beyond chance rate well before they can explicitly formulate one or more of the underlying rules of one of both grammar sets. The implicit learning hypothesis is thereby confirmed. Recently, doubts have been formulated about the sensitivity of the measurement to assess the transition from pre-conceptual to conceptual phase or in other words if, what has been labeled as implicit knowledge, maybe was explicit but wasn't expressed as such (Maia & McClelland, 2004). In our confirmatory experiment we therefore introduced a potential extra reward for a number of subjects (see appendix 1 for details) rather than a simple question to assess if the subjects had any idea about the rules underlying the grammars. This more sensitive measure yielded slightly *less* knowledge than the simple question. We therefore assume that indeed the assessment of the moment knowledge becomes explicit is valid. Thus implicit learning has been shown with this modified Damasio paradigm.

The somatic marker hypothesis is also confirmed in this experiment. Skin conductance responses were significantly larger before making incorrect decisions than before making correct decisions in the pre-conceptual period. A similar finding occurs with the pupil size. In both experiments we have measured small but consistent differences in pupil size when looking at incorrect alternatives compared with the pupil size when looking at the correct alternative. This effect was consistenly largest in the pre-conceptual period. The general somatic marker hypothesis is therefore replicated with a task that has more ecological validity than the original gambling task. The results of the detailed trial by trial analyses in search for causal relations between the somatic marker and the final choice gives a bit confusing results. In the pilot experiment the results seemed to indicate that subjects in general did pick the alternative that generated the smallest pupil size. In the formal experiment the effect sizes for this effect were much smaller. There may be many reasons for this discrepancy. In the Pilot experiment much of the data preprocessing was done manually while in the formal experiment this was replaced by automatic and objective methods. Amongst others this resulted in also taking into account responses where subjects looked shorter than 0.66 seconds. It takes time for the pupil size to adjust and maybe the allowance of these short reponses did result in wrong classification for which alternative resulted in the largest response. (see also the last paragraph of this section)

However in pilot as well as formal experiment the supporting role of somatic Is strongest in the pre-conceptual period. This indicates that the somatic marker process plays its part only in situations where problems are under-specified and a complete analytical solution is therefore not possible. In contrast, conscious decision-making does not seem to be influenced by emotional, physiological bodily reactions. When people become aware of the best options at hand we do no more generate anticipatory somatic markers because we KNOW what to do, so what is the use of listening to our body. From that moment on, cognitive processes are overruling the decision-making process and the advisory role of emotions is pushed to the background.

The intuitive decision can now be modeled as consisting out of several processes. In the early trials subjects seem to learn implicitly about features typical for words of planet X or planet Y. The first process is therefore ‘implicit learning’. In many subjects the positive or negative feedback seems result in the association of a emotional component with the implicit knowledge. This is the process of somatic marking. It runs more or less in parallel with the implicit learning processes. And then finally some of the subjects ‘use’ the somatic marker in the decision process. They tend to listen to what this bodily response tells them. It is especially the last step that might be trained by bio feedback methods. Future research should focus on individual differences and on the possibility to train people to increase their sensitivity to the somatic marker.

The reported results show clearly that the new paradigm enables one to get a much better insight in the role of the somatic marker in the decision process than was possible with the measurement of skin conductance during a gambling task. In this new approach, early reflections about the alternatives and the effect thereof on the somatic marker are separated from final decisions. These early results support the view that people actually USE their somatic markers in the decision-making process although without further detailed analyses it is unclear how exactly this relation between somatic marker and choice is established.

Although mean pupil size is the commonly used dependent variable when trying to measure cognitive load or emotional valence etc., the rate of pupil size change starting within 100 msec after the moment of fixation on a word is highly correlated and is much less affected by the systematic errors a.o. introduced by camera displacement.
Therefore pupil size change rate has to be explored as a dependent variable in future research. Using pupil size change rate has also the advantage that results of mean pupil size are contaminated by the duration the subject is looking.

References

Beatty, J. Task-evoked pupillary responses, processsing load, and the structure of processsing resources. *Psychological Bulletin* **91,** 276-292 (1982).

Bechara, H. Damasio, A. Damasio A. R. and Andersen, S. W. Insensitivity to future consequences following damages to human prefrontal cortex. *Cognition* **50,** 7-15 (1994).

Bechara, H. Damasio, A. Tranel, D. and Damasio A. R. Deciding advantageous before knowing the advantageous strategy. *Science* **275,** 1293-1295 (1997).

Bechara, A. Damasio, H and Damasio, A. R. Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex* **10,** 295-307 (2000).

Bechara, A. Dolan, S. Denburg, N. Hindes, A. Andersen, S. W. and Nathan, P. E. decision-making deficits, linked to dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia* **39,** 376-389 (2001).

Bechara, H. and Damasio, A. Decision-making and addiction (part 1): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia* **40,** 1675-1689 (2002).

Bechara, A, Dolan S and Hides, N. Decision-making and addiction (part 2): Myopia for the future or hypersensitivity for reward. *Neuropsychologia* **40,** 1690-1705 (2002).

Bierman, D. J. and Cleeremans, A. The use of an implicit grammer task for the study of the somatic marker hypothesis. *Draft April 30, 2002.*

Bradley, M. M. and Lang, P. J. Affective reactions to acoustic stimuli. *Psychophysiology* **37,** 204-215 (2000).

Cleeremans, A. Destrebecqz, A. and Boyer, M. Implicit learning: news from the front. Trends in Cognitive Sciences **10**, 406-416 (1998).

Damasio, A. R. Descartes error. Emotion, Reason and the human brain. Grosset/Putman, NewYork (1994).

Damasio, A. R. Fundamental feelings. *Nature* **413,** 781 (2001).

Filipic, S. The origin of somatic markers: A suggestion to Damasio's theory inspired by Dewey's *Ethics*.

Hess, E. H. Pupillometrics. In:Greenfield, N. S., Sternbach, R. A. (Eds.). Handbook of Psychophysiology. New York, pp. 491-531 (1972).

Janisse, M. P. Pupil size, affect and exposure frequency. *Social Behavior and Personality* **2,** 125- 146 (1974).

Lang, P. J. The emotion probe: Studies of motivation and attention. *American Psychologist* **50,** 372-385 (1995).

LeDoux, J. The Emotional Brain: The Mysterius underpinnings of emotional life. Touchstone. New York (1998).

London, E. D. Ernst, M. Grant, S. Bonson, K. and Weinstein, A. Orbitofrontal cortex and human drug abuse: functional imaging. *Cerebral Cortex* **10,** 334-342 (2000).

Maia, T.V and McClelland J.L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *In Press*

Manes, F. sahakian, B. Clark, L. Rogers, R. Antoun, N. Aitken, M. and Robbins, T. Decision-making processses following damages to the prefrontal cortex. *Brain* **125,** 624-639 (2002).

Murphy, S. T. and Zajonc, R. B. Affect, Cognition, and Awareness: Affective Priming With Optimal and Suboptimal Stimulus Exposures, *Journal of Personality and Social Psychology,* **64,** *723-739* (1993).

Partala, T. and Surakka, V. Pupil size variations as an indication of affective processsing. *International Journal of Human-Computer Studies* **59,** 185-198 (2003).

Reber, A. S. Implicit learning of artificial grammars. *Journal of verbal learning and verbal behavior* **6,** 855-863 (1967).

Siegle, G. J. Granholm, E. Ingram, R. E. and Matt, G. E. Pupillary and reaction time measures of sustained processsing of negative information in depression. *Society of Biological Psychiatry* **49,** 624-636 (2001).

Siegle, G. J. Steinhouer, S. R. Thase, M. E. Stenger, V. A. and Carter, C. S. Can't shake that feeling: event-related fMRI assesment of sustained amygdala activity in response to emotional information in depressed individuals. *Society of Biological Psychiatry* **51,** 693-707 (2002).

Volkow, N. D. and Fowler, J. S. Addiction, a disease of compusion and drive, involvement of the orbitofrontal cortex. *Cerebral Cortex* **10,** 318-325 (2000).

APPENDIX 1
Two methods for knowledge elicitation

FB1:
After each 10 trials the following question:
“On what do you base your choices”
If answer is unclear:
“What do you mean by that”
If a rule is mentioned:
“When do you think you discovered this way to make a choice?”

FB2:
As FB1, but in addition to that
After 20 trials (if no correct rule has been mentioned):
If you can formulate a correct rule now we increase your score with 100 Euro. I cannot tell you if the rule might be correct or not, after the session the final score will be calculated and thus by answering a correct rule now you increase your chances to win a book credit (of 50 Euro).

APPENDIX 2
Instruction

Which word comes from planet X?
You will see two words on the screen. One word comes some Planet X and the other of some planet Y.
The goal for you is to indicate which of those two words comes from the planet X.
After each choice you will get feedback on the screen. In the beginning you cannot know which of the two words is the one for X so in that case just follow your feelings.
There are in total 100 trials. After each 10 of such trials we will ask you on how you come to a decision.

Each time you correctly identifies the word from planet X your score will be increased. In case of an incorrect choice the score will be decreased. The person who gets the highest score will receive a book credit of 50 euro.

The experimenter will explain how you can enter your choices into the computer. If you still have questions you might ask the experimenter.

Good Luck!

APPENDIX 3

Informed consent

In this experiment two ‘words’ are presented on the screen. These words are composed of strange characters like #,], * and +. You have to make a choice between the two words by means of the button box. After each 10 trials you will be asked by the experimenter to explain how you decide.

During the experiment we make a video of your eye. This will only work if you do not move. That’s why we need this chair with the head mount to fixate your head. We will fixate with ‘klittenband’. This is a commonly used method but it could be experienced as a bit uncomfortable. You will have to sit like that for about 20 minutes.

The camera just registers your eye. The experimenters will not use any actual image but use the measurements of your pupil dilation that are calculated by the computer. There is an infrared light source. It’s intensity is much lower than what is generally considered to be acceptable so there is no risk for any harm.

It is always possible to quit during the experiment. Your participation is voluntary.

We would like to ask your permission for the use of the data. All results are anonymous, nobody except the experimenters will have access to the data.

If you still have questions you can contact any of the following persons: Eva Lobach, Jenneke van Ditzhuyzen or Dick Bierman

---

I hereby declare that I read and understood the text above and give permission for the use of the data that are registered in this experiment.

Name:
Amsterdam, Date
Signature:

APPENDIX 4
Exit Interview

1. Did you find this task difficult?

2. Could you see the words well?

3. How seriously did you participate?

4. Can you tell us a bit more about your strategy?

5. Did you for whatever reason sometimes gamble or did you hide for us a rule that you had discovered (for instance because you weren't very sure about the rule)?

6. Are there other things that might be relevant? (Like you are very tired or similar)

7. What do you think was the goal of this experiment?

8. Do you have any further questions?