

BACHELOR THESIS IN PSYCHOLOGY

*Autumn 2007*

Department of Behavioural Sciences

# Valence asymmetry in impression formation

Author  
Montathar Faraon

Supervisor  
Georg Stenberg

## *Valence asymmetry in impression formation*

*Are implicit attitudes as easily formed as explicit attitudes? Fifty-four participants carried out an experiment regarding the behaviors of a fictional protagonist. Attitudes were obtained by the Implicit Association Test (IAT) and by explicit evaluations. The results in the current study showed tendency towards a significant persistence effect for explicit attitudes and a positivity bias effect for implicit attitudes.*

*Keywords:* asymmetry, attitudes, impression formation, bias, implicit, IAT

During a long period of time scientists have been dealing with good and evil as concepts related to the reflection of reality and meaning. These reflections have given birth to new philosophical and religious systems that could be considered among the most remarkable achievements done by the human species. These achievements would not have been possible to discover without our extraordinary information-processing device, the brain (Peeters, 1986). For centuries a growing catalogue of biases and asymmetries have pointed out that negative information influences people's evaluation system in the brain more strongly than comparably extreme positive information. One area where this is especially evident is impression formation (Peeters & Czapinski, 1990; Skowronski & Carlston, 1989). To understand impression formation, one has to investigate how the reasoning systems work in the brain.

In an illustrative study, Sloman (1996) argued that there are two different and independent systems of reasoning in our brain with regards to what information they process and how they process it. The first system is called the *slow-learning system* and uses interconnected associations in memory that are based on similarity and contiguity. This system is characterized by associations in memory that are developed and formed over time. The second system is called the *fast-learning system* and is based on logic and symbolic representations on a moderately high level of cognitive processing. An overview of these systems is presented in Table 1.

**Table 1.** *An overview of the different systems of reasoning.*

<i>Characteristic</i>	<i>Slow-learning system</i>	<i>Fast-learning system</i>
Principles of operation	Similarity	Symbol manipulation
Source of knowledge	Personal experience	Language and culture
Relations	Associations	Logical
Cognitive functions	Fantasy and creativity	Explanation

According to Smith and DeCoster (2000) implicit attitudes are assumed to follow the basic principles of similarity and association. Implicit attitudes can be described as preferences that exist outside of conscious awareness or conscious control. Thus, giving us the possibility to understand, from the slow-learning system as point of departure, how the implicit attitudes are formed and function. In contrast, the fast-learning system is nicely compatible with explicit attitudes, indicating that people can occasionally control their expression of explicit attitudes.

Indeed, empirical studies have shown that implicit and explicit attitudes predict different kinds of behaviour, such as spontaneous and nonverbal or deliberate and self-presentational (Jellison, McConnell, & Gabriel, 2004; McConnell & Leibold, 2001).

To measure these attitudes, one may use explicit evaluations for the explicit attitudes and the Implicit Association Test (IAT) for the implicit attitudes. According to Greenwald, McGhee, & Schwartz (1998) the measurements in the IAT are defined in terms of association strengths. This means that the respondents respond faster when the concept and attribute are strongly associated (e.g., *flower* and *pleasant*) than when they are weakly associated (e.g., *insect* and *pleasant*).

On the basis of the discussed information, one may broadly argue that the slow-learning system and the fast-learning system form attitudes in different ways. In other terms, are implicit attitudes as easily formed as explicit attitudes? Are we more influenced by negative information than positive? Do we have a preconceived opinion – a fundamental base of values – about other people? Do men and women form attitudes differently from each other? Hence, the purpose of the current study is to investigate asymmetry of valence in attitude formation with certain consideration for implicit and explicit attitudes.

### ***Previous research***

Despite the central role the concept of attitudes has occupied in social psychology, relatively little focus has been given to the processes involving attitude formation per se. Eagly and Chaiken (1993) concentrated in a study on questions regarding attitude change, structure and function but also influences of attitudes. In the same study, the authors pointed out that future directions should investigate attitude formation. The authors referred that the current field had a “lack of attention to the developmental issue of how attitudes are formed and become strong” (p. 681).

Similarly, Fazio, Eiser, and Shook (2004) described in their study that there was currently a lack of research in experimental social psychology that examined the formation of attitudes toward novel objects, that is to say ones for which individuals have no relevant a priori knowledge. More specifically, the authors expressed a sense of criticism toward the literature that lacked in describing systematic research examining how attitudes toward novel objects develop over time as a function of individuals’ own exploratory behaviour. The author pursued their criticism and commenced an exploratory investigation based on a computer game that involved participants to spend time in a virtual world. Because the world consisted only of beans the game came to be called “BeanFest”. The participants’ goal was to survive in the virtual world by avoiding a decrease of one’s energy level. The beans themselves varied in the outcomes they produced. Some provided energy, whereas other depleted energy. During the course of the game the authors were in a position to ask whether any developed attitudes would generalize to novel stimuli. The results showed that participants improved their learning performance over time. This was examined by analyzing the phi coefficient, which relatively increased for each of the learning phases. Furthermore, the results indicated that positive beans were significantly less likely to be classified correctly than the negative beans; indicating an apparent asymmetry in learning.

A study by Rydell, McConnell, Mackie and Strain (2006) investigated how implicit and explicit attitudes differ in sensitivity with consideration to processing different kind of information. More specifically, the authors of this study investigated the sensitivity to subliminally presented primes that provide valence target-evaluation as associations but also consciously accessible information that requires cognitive processing. The results showed that implicit and explicit attitudes are formed and revised on the basis of different types of information with consideration to our underlying systems of reasoning.

In contrast, Peeters and Czapinski (1990) examined the negativity bias, which can be defined as the greater impact negative information has on a subject than equally intense positive stimuli. The point of departure in this study was related to two manifestations of the phenomenon: (1) potential costs weigh more heavily than potential gains in making decisions under risk, and (2) evaluative negative information weighs more heavily than evaluative positive information in the formation of evaluations. The outcome of the study revealed that more weight was accorded to potential losses than to potential gains in the decision process. Furthermore, the results showed that negative information has a stronger impact on our mind than positive information in attitude formation.

Moreover, Peeters (1986) discussed how the positivity offset and the negativity bias are coherent with the concepts of 'good' and 'evil' as categories of implicit knowledge. In this manner the author found that verbal descriptions of negative states could take prominence over positive information. This could develop positive states to turn into bad ones by one single negative term denying the evoked positive states. For example, the terms 'extraordinary, teacher, best, student and daylight' may all elicit positive associations, but in conjunction with one single negative term such as 'murders' they may develop a dreadful message such as: 'An extraordinary teacher murders his best student in broad daylight'.

Ito, Larsen, Smith and Cacioppo (1998) showed the same effects in evaluative reactions to words in a dataset called the Affective Norms for English words. In the same study, the authors investigated how the positivity offset and negativity bias could describe the functional properties of the underlying positive and negative motivational systems. Since variation is the engine of natural selection, the authors could expect that adaptive benefits would vary among individuals. Hence, individuals with a stronger positivity offset would respond more at low levels of activation while individuals with a stronger negativity bias would show stronger responses to negative evaluative information. The results in the first study revealed a negativity bias with a significantly steeper slope than the slope for positivity in both of the conducted sessions. Surprisingly, the second study showed a main effect of gender for positivity offset with women showing a stronger positivity offset but also a main effect showing that women had a stronger negativity bias.

Also related to this issue is how the asymmetry between the positivity offset and negativity bias can be explained through systems of reasoning. In an earlier study, Rydell and McConnell (2006) proposed that explicit attitudes form and change through the use of fast-learning, rule-based system, whereas implicit attitudes form and change through the use of slow-learning, associative system. Their work was aimed to understand whether slow-learning and fast-learning systems could account for implicit and explicit attitudes. In order to examine this approach the authors gave participants information about a novel target person (*Bob*) in a learning paradigm that presented neutral (e.g., Bob likes all different kinds of music), positive (e.g., Bob donates his old clothes to the needy) and negative (e.g., Bob stole coffee from a gas station) behavioural information. The results showed that implicit and explicit attitudes changed because they are governed by different systems of reasoning. Indeed, implicit attitudes formed according to the slow-learning, associative system, whereas explicit attitudes formed in coherence with the fast-learning, rule-based system. At the same time the results indicated that feelings from the subliminal presentation did not have a substantial affection on explicit attitudes.

More broadly, this study showed support for the two-attitudes model's suppositions; (1) People can have different implicit attitudes and explicit attitudes about an attitude object while at the same time (2) implicit attitudes are not changed equally as explicit attitudes in terms of time. These two suppositions are important to understand the current research and models of attitude change (Wilson, Lindsey, & Schooler, 2000).

Most models on attitude change are based on the dual-attitude framework, that is to say automatic and controlled processes. Two well-known models in this framework are the equation likelihood model (ELM) and the motivation and opportunity to deliberate model (MODE). The former predicts according to Petty and Wegener (1998) that once an attitude is changed by the influence of compelling arguments or by the sight of persuasion (e.g., attractiveness), the original attitude state will no longer exist. Proponents of the latter model argue that differences between implicit and explicit attitudes are clear signs that people differ in the ability to modify the initial activation of an attitude in memory (Fazio, 1995).

In spite of this, Cunningham and Zelazo (2007) have pointed out in their study that dual-attitude models are too simple to give a clear picture of the dynamic way in which affective attitudes and reflective processing interact. They proposed a new model in their study, namely the iterative-reprocessing (IR) model which in short highlights the reprocessing of evaluative information about valenced stimuli but also the interaction among automatic and reflective processes. The model involves neural networks that are hierarchically arranged in a way that common processes become involved in generating evaluations throughout a cycle. The results of their neurological experiment showed that information is processed dynamically through the brain by using recursive feedback loops. Conclusively, this study showed a dynamic view of information processing in contrast to a dual-attitudes perspective, which is generally based on a relatively static thought. Further research in this area would emerge a clearer picture of how attitudes arise from cognitive processes.

Another important implication, noted by Cacioppo, Gardner and Berntson (1997), was that the positivity offset and negativity bias are posited as having separable neurophysiological substrates that have different functional outputs. This can often be seen in risk-taking research, where losses loom larger than gains.

Taylor (1991) showed that an organism's physiological responses become mobilized as a result of a negative event. This is particularly interesting in the sense that evolution has taught the human species to act in such way. This reaction could be described as learned evaluative responses that are acquired through the principles of learning. The degree of evaluation acquired by the attitude object is, according to Kerpelman and Himmelfarb (1971), a function of the consistency of reinforcement or strength of association. The predicted effects of partial reinforcement have been well documented in classical conditioning but also in concept-formation. However, the inverse relationship between consistency of reinforcement and counterconditioning is not yet substantiated in a clear way. Thus, Kerpelman and Himmelfarb (1971) investigated this issue based on the concept-formation paradigm introduced by Rhine and Silun (1958). The experiment presented different evaluative traits in which participants would indicate whether or not the trait is characteristic of the people. In order to assess the effects of counterconditioning the degree of association was manipulated by varying the proportion of traits that are reinforced as "characteristic". The results showed that the greater the strength of association, the more likely was the subject to say that a positive trait was associated instead of a negative trait with the attitude object. Furthermore, the results suggested that the resistance to attitude change could be produced by a limited amount of partial reinforcement.

Reinforcement in attitude learning is a very important area to explore because a recent study by Shook, Fazio and Vasey (2007) investigated to which extent a learning bias in attitude formation could be connected to emotional disorder symptoms. To examine this relationship, the authors used the computer game called BeanFest together with three questionnaires concerning emotional disorder symptoms, CSQ (Cognitive Style Questionnaire), BDI (Beck Depression Inventory II) and BAI (Beck Anxiety Inventory). Overall, the results showed that learning of the game correlated with the scales such that

poorer learning was associated with a more negative cognitive style and a tendency toward greater anxiety and vice versa.

Todorov and Willis (2006) investigated another interesting perspective in attitude formation, namely the minimal conditions under which people make trait inferences from the facial appearance of other people. In all the experiments, faces unfamiliar to the participants were shown during 100 ms, 500 ms and 1000 ms. The results revealed that a minimal exposure time as a tenth of a second was sufficient in order for people to make a specific trait inference from a facial appearance.

Based on previous theoretical knowledge and research results the present study aims to investigate several hypotheses which are as follows: (1) Explicit attitudes will have a positive bias in the beginning but change strongly for incongruous negative information; (2) Negative information will be more diagnostic than positive; (3) Implicit attitudes will not show any asymmetry of valence.

## **Method**

### *Participants*

Fifty-four participants (31 male and 23 female) aged between 17 and 51 years from Kristianstad University and Finnvedens upper secondary school volunteered to complete a computer-based experiment. Participants were randomly assigned to one of two conditions (positive vs. negative) in which they were all provided with neutral information in the beginning. They were then exposed to either positive or negative information. Finally they received once again neutral information. Their reactions were measured on three occasions, producing a 2 x 3 mixed design (one between-subjects factor and one within-subjects factor) for each of the two dependent variables: implicit and explicit attitudes.

### *Materials*

To assess explicit attitudes, participants judged on three occasions during the computer-based experiment how sympathetic the protagonist *Oscar* was on a scale ranging from 1 (very unsympathetic) to 9 (very sympathetic). The ratings of likeability on a scale from 1 to 9 were transformed to a scale ranging from -4 to +4 to facilitate comparisons with the other measures.

The Implicit Association Test (IAT) was used to assess participants' implicit attitudes. According to (Greenwald, McGhee, & Schwartz, 1998; Greenwald, Nosek, & Banaji, 2003) the measurements in the IAT are defined in terms of association strengths. In the IAT participants respond to a series of items that are to be classified into four categories, two representing a concept discrimination such as police versus criminals and two representing an attribute discrimination such as moral versus immoral. The IAT measures are interpreted in terms of association strengths by assuming that participants respond more rapidly when the concept and attribute mapped onto the same response are strongly associated (e.g., police and moral) than when they are weakly associated (e.g., criminals and moral). The concept discrimination pair in the current study was a protagonist called Oscar and two male figures as non-Oscar and the attribute discrimination pair was positive/negative adjectives.

This measure was computed as  $d$ , the parameter recommended by Greenwald, McGhee, & Schwartz (1998). It is an effect size type of measure with ratio scale properties, i.e. it has a natural zero, indicating neutral attitude. Positive numbers indicate a positive attitude to the protagonist (i.e. greater ease in associating Oscar with favourable qualities). Negative

numbers indicate a negative attitude. The IAT consisted of 42 stimuli<sup>1</sup>: 6 pictures of Oscar, 16 pictures of White men who were not Oscar, 10 positive adjectives (e.g., *superb*), and 10 negative adjectives (e.g., *horrible*). The experiment was a modified version from previous research by Rydell and McConnell (2006), featuring 11 blocks with a total of 308 trials<sup>2</sup>.

The two measures mentioned formed the *original* data set. Additionally, a rectified data set was calculated that consisted of the original data for the positive group, but for the negative group, there was a change of sign. This is the practice followed by Rydell & McConnell (2006). By giving effects in the two groups the same direction, facilitated comparisons between their sizes could be made. In the following, the original data are presented in figures, but – unless explicitly stated otherwise – statistical tests are performed on the rectified data set.

The hypotheses in the current study were tested by a priori contrasts, formed out of measurements on three critical occasions: at the end of the first neutral block, at the end of the positive/negative attitude formation block, and at the end of the last neutral block. Contrasts were of three kinds: positivity bias effects, attitude formation effects and persistence effects. The positivity bias hypothesis was tested by measurements on the first occasion, using the original data set. Because the participants had been exposed only to an equilibrium of positive and negative information up to that point, any deviation from zero can be taken to mean a positivity bias. Attitude formation effects were measured by the difference between the second and first measurement occasions, i.e. change scores. These scores were tested both in a within-subjects test and a between-subjects test. The former tested whether our manipulation had any overall effect, the latter whether it was larger for those exposed to predominantly negative information. Persistence effects were measured by the difference between occasion three and occasion one. If attitudes were brief and evanescent, this difference would be zero. If, however, attitudes were integrated over time, the effects of earlier information would linger on and affect the last measurement occasion, preventing a return to zero.

### *Procedure*

Participants were seated individually in front of a computer during the entire experiment. In the beginning of the experiment, participants were given introductory instructions on how to complete the experiment by using the left and right keys on the keyboard. In Block 1, participants were presented with 20 behaviors of the protagonist Oscar that were either positive or negative in valence. After each behavior, participants indicated whether they believed that the behavior (e.g., Oscar stole money and jewelry from relatives he was living with) was characteristic or uncharacteristic of Oscar by pressing the left-arrow (characteristic) or the right-arrow (uncharacteristic). The feedback in behavior blocks consisted of a blue text message (you are correct) or a red text message (you are incorrect). The main experimental manipulation was the different proportion of positive and negative feedback given between the positive and negative group according to Table 2.

<sup>1</sup> One of three male photographs was randomly presented as “Oscar” while the other two male photographs were presented as “non-Oscar”. The photographs did not produce any effects. The positive and negative behaviors were developed by McConnell, Sherman, & Hamilton (1994); they were presented only once and were selected to ensure that they were not contradictory.

<sup>2</sup> All trials in the critical blocks were retained for analysis. Responses faster than 400 ms or slower than 10,000 ms were excluded. According to recommendations by Greenwald, Nosek, & Banaji (2003), error responses were recalculated as mean value of correct response plus penalty of 600 ms.

**Table 2.** Behaviors and correct response balance expressing the experimental manipulation.

<i>Behaviors / Correct response</i>	<i>Occasion</i>		
	<i>1</i>	<i>2</i>	<i>3</i>
Positive / Characteristic	5	18 / 2	10
Positive / Uncharacteristic	5	2 / 18	10
Negative / Uncharacteristic	5	2 / 18	10
Negative / Characteristic	5	18 / 2	10

**Note.** Feedback differed in the second part between groups. The first number applies to the positive group, the second to the negative one. Through the three blocks, the positive group received 50, 90 and 50 % positive information about the protagonist. The negative group received 50, 10 and 50 % positive information.

Thereafter, participants completed two training blocks where they were told that they would be receiving information about a protagonist called Oscar. During Block 2, participants received 20 photos of Oscar and non-Oscar and in Block 3 they were presented with 20 positive and negative adjectives. In Block 4 and 5 (Combination 1), participants judged 56 combined trials that were “Oscar and positive adjectives”, “Oscar and negative adjectives”, “non-Oscar and positive adjectives” or “non-Oscar or negative adjectives”. The difference between the two blocks was that the assignment of response keys assigned to the attribute discrimination pair was reversed. Block 6 consisted of the main experimental manipulation where participants performed the same judgment as Block 1 except that 40 positive or negative feedbacks were given. In Block 7 and 8 (Combination 2), participants’ performed the same judgment task as in Block 4 and 5 (Combination 1). In Block 9, participants performed the same judgment task as Block 1 and 6 except that they were given 40 neutral feedbacks. Finally, in Block 10 and 11 (Combination 3), participants’ performed the same judgment task as in Combination 1 and 2. The experiment task sequence is presented in Table 3.

**Table 3.** The experiment task sequence used to assess implicit and explicit attitudes.

<i>Block</i>	<i>Task</i>	<i>Trials</i>	<i>Information</i>
1	Behavior learning	20	Neutral
2	Facial recognition training	20	Neutral
3	Positive/negative adjectives training	20	Neutral
4-5	IAT (Combination 1)	56	Neutral
6	Behavior learning	40	Positive / Negative
7-8	IAT (Combination 2)	56	Positive / Negative
9	Behavior learning	40	Neutral
10-11	IAT (Combination 3)	56	Neutral

**Note.** A total of 11 blocks featuring 308 trials were conducted throughout the experiment.

After each response, participants were given feedback whether the indicated answer was correct or not for 2 s. The feedback in IAT blocks consisted of a blue plus sign (correct) or a red cross (incorrect) positioned in the centre of the computer screen. The ordering of the combinations was counterbalanced in accordance with the experimental condition, the participant number.

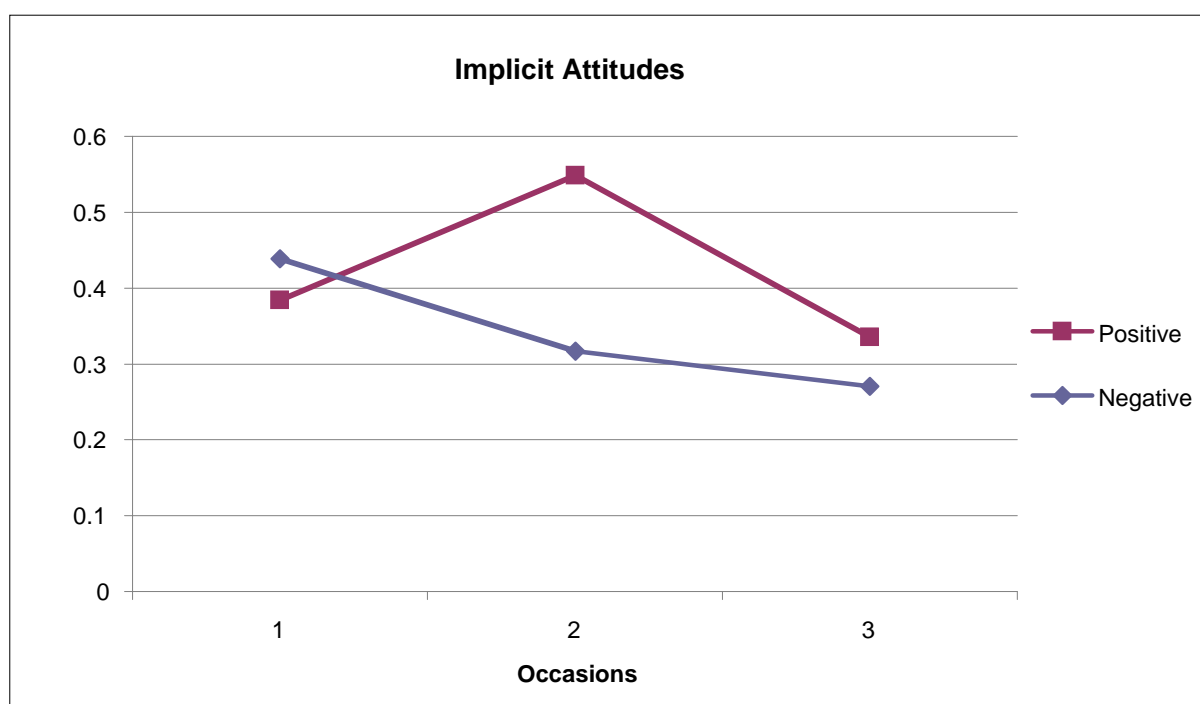


## Results

Concerning the question whether implicit attitudes are as easily formed as explicit attitudes, an attitude formation effect was calculated by subtracting the result of block one from the result of block two. A  $t$  test for the original data set showed a significant within-subjects attitude formation effect for participants' implicit attitudes,  $t(53) = 2.160$ ;  $p = .035$ .

Further, analysis of variance (ANOVA) of the original data set revealed an interaction effect between the occasions and the condition groups for the implicit attitudes,  $F(1, 51) = 5.11$ ,  $p = .028$ ,  $\eta^2 = .091$ . These results showed that the condition groups formed implicit attitudes according with the valence of presented information on the second occasion. These results are illustrated according in Figure 1.

Additionally, analysis of variance (ANOVA) of the rectified data revealed a between-subjects main effect for the condition groups,  $F(1, 51) = 71.05$ ,  $p < .001$ ,  $\eta^2 = .582$ .

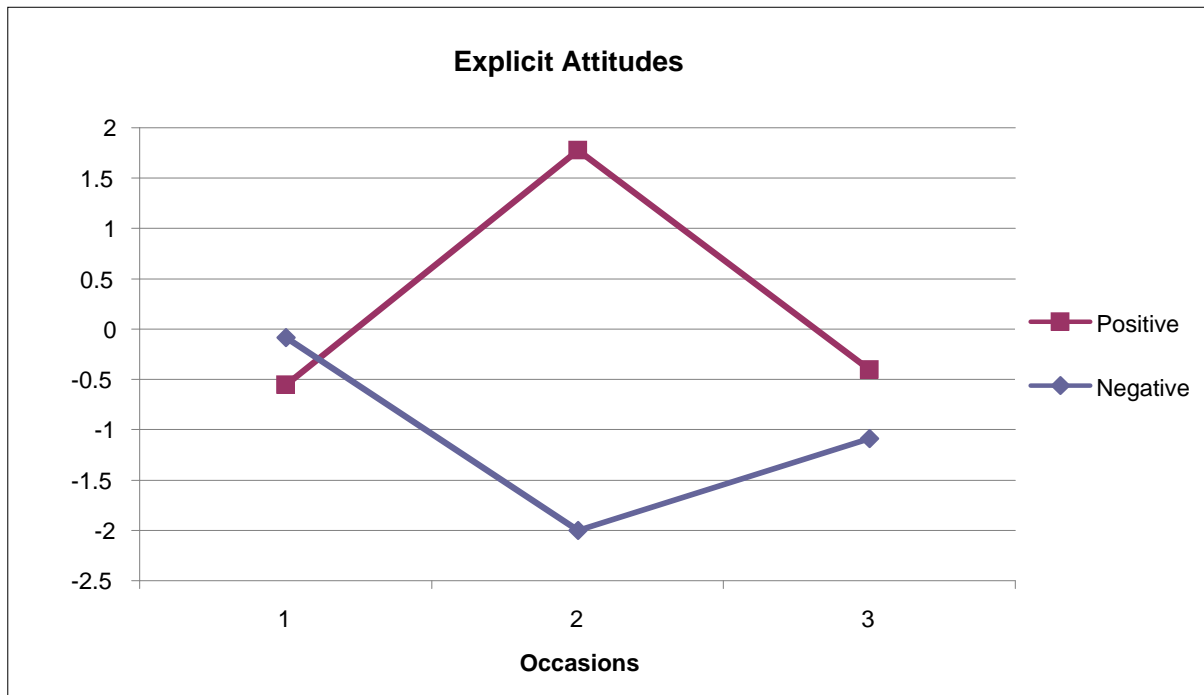


**Figure 1.** Mean values for the implicit attitudes in the two condition groups.

In order to investigate whether participants had a preconceived opinion about other people, one would expect a positivity bias effect. The results for the positivity bias effect were calculated by using original data in the analysis of the first block. The analysis showed a significant positivity bias effect for participants' implicit attitudes,  $t(53) = 5.612$ ;  $p < .001$ , and a weak tendency for explicit attitudes,  $t(53) = 1.704$ ;  $p = .094$ .

Further, analysis of variance (ANOVA) of the original data set revealed an interaction effect between the occasions and the condition groups for the explicit attitudes,  $F(1, 51) = 57.99$ ,  $p < .001$ ,  $\eta^2 = .532$ . Participants' explicit attitudes formed in response to the valence of written behavioral information.

Additionally, the results for participants' explicit attitudes showed a within-subjects persistence effect,  $t(53) = 2.369$ ;  $p = .021$ , and a tendency towards significance for the between-subjects results,  $t(52) = 1.797$ ;  $p = .078$ . Further, analysis of variance (ANOVA) of the original data set revealed a between-subjects main effect for the condition groups,  $F(1, 51) = 33.26$ ,  $p < .001$ ,  $\eta^2 = .395$ . These results are illustrated in Figure 2.



**Figure 2.** Mean values for the explicit attitudes in the two condition groups.

Regarding the question whether negative information has a stronger influence than positive, one would expect an attitude formation effect in a between-subjects test. This effect was calculated by subtracting rectified data of block one from block two. The results showed no significant attitude formation effect for implicit or explicit attitudes. No gender differences were obtained for any of the three effects in any of the dependent variables. Thus, men and women did not differ in attitude formation according to the findings in this study.

## **Discussion**

The main goal of the present study was to investigate asymmetry in processing positive and negative information. The results in this study revealed several important findings that support prior research in attitude formation (Ito, Larsen, Smith, & Cacioppo, 1998; Peeters & Czapinski, 1990) but was inconsistent with other attitude formation research (Rydell & McConnell, 2006; Rydell, McConnell, Mackie, & Strain, 2006; Sloman, 1996; Smith & DeCoster, 2000).

Sloman (1996) and Smith & DeCoster (2000) argued that implicit and explicit attitudes are formed consistent with different underlying systems of reasoning, associative and rule-based system. According to Sloman (1996) and Smith & DeCoster (2000) these two systems function as two computational resources that work cooperatively with each other. Although each system may be able to compute a response for a certain problem, one may not always find the responses to be consistent. Our awareness of an inference may play a major role in the outcome of our thoughts. Consider for instance the anagram *resovle* for which the answer most likely become implicitly available, i.e. resolve. In this case, one is only conscious about the result, not the process that leads to the result. Furthermore, consider a more difficult anagram such as *yvolnutarin* for which the correct answer most likely becomes explicitly available through the rule-based system (e.g., reposition of different letters), which consequently made you both conscious about the process and the result.

Similarly, Rydell & McConnell (2006) claimed that explicit attitudes were changed dramatically unlike implicit attitudes by the introduction of counterattitudinal information. Concerning implicit attitudes, if they were based on the slow-learning system, then they would be slower to change and simultaneously time-consuming to lose. However, explicit attitudes showed a tendency towards a significant persistence effect, which could be given a plausible explanation; implicit attitudes are not necessarily slower than explicit attitudes, but rather less sensitive. This could be examined by reproducing the experiment in the current study with higher time resolution and more measurements. In this manner, one has to take under consideration that the important factor is the course of time and the speed of information processing. There might have been a high level of probability of obtaining a significant persistence effect for the explicit attitudes if the number of participants were higher. Consequently, the findings in the current study showed no support for a correspondence between implicit attitudes and the associative system (slow-learning) on the one hand and explicit attitudes and the rule-based system (fast-learning) on the other.

Moreover, the results of the cognitive evaluation showed that participants found negative information to be more diagnostic than positive. These findings support prior research by Ito, Larsen, Smith, & Cacioppo (1998) and Peeters & Czapinski (1990) where the former revealed a negativity bias with a significantly steeper slope than the slope for positivity while the latter found evidence supporting the idea that negative stimuli hold stronger informational value than positive ones. Ito, Larsen, Smith, & Cacioppo (1998) found that women had both a stronger positivity offset and a stronger negativity bias than men. Concerning the positivity offset, women showed a greater tendency toward positive content than equally negative content at a low level of activation. The negativity bias showed that women had a greater change in output in the negative motivational system than in the positive motivational system. One might have expected to find similar differences between men and women in the current study, but the results revealed no gender differences in attitude formation.

Maybe the most important and difficult question to answer in the current study concerns the positivity bias that was revealed for participants' implicit attitudes in the first IAT block. What accounted for this positivity bias? A plausible explanation for this issue is referred to by Skowronski & Carlston (1989) and Cacioppo, Gardner, & Berntson (1997) as the positivity bias or the positivity offset. This bias is often found when participants perceive positive cues as more diagnostic than negative ones. This might have been elicited because the participants perceived Oscar much more during the initial blocks than non-Oscar. Future directions should follow up this issue with another experiment. What would happen if one should use a more recognized facial object as non-Oscar, such as a famous actor? Will the positivity bias for Oscar disappear due to the fact that the actor is easier to connect with positive words? More importantly, what would happen if one would use a more negatively loaded person as non-Oscar, such as a famous murderer?

Regarding the methodology of this study, standardized attitude measures were adapted from previous research by Rydell & McConnell (2006) in terms of validity and reliability. Although one may be critical to the fact that the experiment was conducted in an artificial environment the content had parallels to the reality, such as media. For example, some participants reported after each session a dissonance regarding how information was presented about the protagonist, from mixed information to either positive or negative to finally mixed. In contrast to media, one day a person might be presented as heroic while the next day as villainous, which is often seen in political contexts. Conclusively, during the experiment session it was noted that a number of participants made some error between the assigned keys. To counteract this, error responses were recalculated according to Greenwald, Nosek, & Banaji (2003) as the mean value of the correct response plus a penalty of 600 ms.

## References

- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review, 1*, 3-25.
- Cunningham, W., & Zelazo, P. D. (2007). Attitudes and evaluation: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences, 11*, 97-104.
- Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.
- Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247-282). Hillsdale, NJ: Erlbaum.
- Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology, 87*, 293-311.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464-1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.
- Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology, 75*, 887-900.
- Jellison, W. A., McConnell, A. R., & Gabriel, S. (2004). Implicit and explicit measures of sexual orientation attitudes: Ingroup preferences and overt behaviors among gay and straight men. *Personality and Social Psychology Bulletin, 30*, 629-642.
- Kerpelman, J. A., & Himmelfarb, S. (1971). Partial reinforcement effects in attitude acquisition and counterconditioning. *Journal of Personality and Social Psychology, 19*, 301-305.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435-442.
- McConnell, A. R., Sherman, S. J., & Hamilton, D. L. (1994). Illusory correlations in the perception of groups: An extension of the distinctiveness-based explanation. *Journal of Personality and Social Psychology, 67*, 414-429.
- Peeters, G. (1986). Good and evil as softwares of the brain: On psychological 'immediates' underlying the metaphysical 'ultimates'. A contribution from cognitive social psychology and semantic differential research. *Ultimate Reality and Meaning. Interdisciplinary Studies in the Philosophy of Understanding, 9*, 210-231.
- Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 1, pp. 33-60). Chichester: Wiley & Sons Ltd.
- Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. Gilbert, S. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (4 ed., pp. 323-390). New York: McGraw-Hill.
- Rhine, R. J., & Silun, B. A. (1958). Acquisition and change of a concept-attitude as a function of consistency of reinforcement. *Journal of Experimental Psychology, 55*, 524-529.

- Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology, 91*, 995-1008.
- Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence inconsistent implicit and explicit attitudes. *Psychological Science, 17*, 954-958.
- Shook, N. J., Fazio, R. H., & Vasey, M. W. (2007). Negativity bias in attitude learning: A possible indicator of vulnerability to emotional disorder? *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 144-155.
- Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*, 131-142.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3-22.
- Smith, E. R., & DeCoster, J. (2000). Dual process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4*, 108-131.
- Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin, 110*, 67-85.
- Todorov, A., & Willis, J. (2006). First impressions: Making up your mind after 100-ms exposure to a face. *Psychological Science, 17*, 592-598.
- Wilson, T. D., Lindsey, S., & Schooler, T. (2000). A model of dual attitudes. *Psychological Review, 107*, 101-126.