Connectionist Model Accounting for Retardation of Cognitive-Dissonance Reduction Caused by Attention-Focus Switching

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Author
Matsumoto, Takao

Publication Date
2012

Peer reviewed

Connectionist Model Accounting for Retardation of Cognitive-Dissonance Reduction Caused by Attention-Focus Switching

Takao Matsumoto (matsumoto@c.dendai.ac.jp)
Tokyo Denki University
5 Senju-Asahi-cho, Adachi-ku, Tokyo 120-8551, Japan

Abstract

A novel connectionist model accounting for cognitive dissonance is described, in which the concepts of self and attention are considered. The model makes it possible to use mathematical formulas to represent the cognitive-dissonance process. Analysis reveals that the model fits experimental data of major paradigms in cognitive dissonance theory and that attention-focus switching causes building-up of cognitive dissonance and retardation of its reduction.

Keywords: cognitive dissonance; connectionist model; attention; self-concept; mathematical analysis

Introduction

Cognitive dissonance theory insists that dissonance is a psychological state of tension that people are motivated to reduce (Festinger, 1957). Dissonance causes feelings of discomfort, unhappiness, or distress. Any two cognitions are dissonant when one of them follows from the obverse of the other. To reduce dissonance, people add consonant cognitions or change evaluations for one or both cognitions to make them more consistent.

Cognitive dissonance theory makes a clear prediction when a firm expectancy is involved as one of the cognitions in question (Aronson, 1969). A well-known example of this is the famous Aesop's fable "The fox and the grapes." In the story, a fox wanted to get some grapes hanging high on vines and leaped with effort, but couldn't get them. Walking away, the fox said, "The grapes are surely sour, and I do not need them." Since the expectation and experience were inconsistent, the fox had cognitive dissonance, which he reduced by convincing himself that the expectation was not appropriate.

Shultz and Lepper (1996) proposed a connectionist model accounting specifically for the mechanism of cognitive dissonance. A constraint satisfaction neural network model was used to simulate data from the several major cognitive dissonance paradigms (Shultz, Leveille, & Lepper, 1999). In it weights between nodes are fixed and activations of units are changed. Dissonance is defined by a formula that is a function of activations of units and weights applied to links in the network. Networks tend to settle into a less dissonant state as activations of units are changed according to update rules. Another connectionist model was proposed by Van Overwalle and colleagues (2002, 2005). They represented attitudes in a feed-forward neural network with the delta-learning rule in which weights are allowed to change. Input nodes represent the features of the environment and two output ports represent behavior and affect. Dissonance is defined as the discrepancy between expected and actual outcomes. They also simulated the experimental results of major cognitive dissonance paradigms. Several other computational models have been reported that deal with attitude phenomena through simulation using constraint-satisfaction or non-constraint-satisfaction networks (Mosler et al., 2001; Petty & Cacioppo, 1986; Read & Miller; 1994; Spellman, Ullman, & Holyoak, 1993).

People are motivated to prioritize to protect their self-system. Self-consistency theory (Aronson, 1969; Thibodeau and Aronson, 1992) emphasizes that self is involved in dissonance arousal and that not only cognitions but also self-concept need to be considered in discussing dissonance. Judgment and assessment for cognitions performed by self possibly become motives for arousal of cognitive dissonance.

On the other hand, attention is an important phenomenon of information processing in cognitive systems (Pisapia, Repovs, & Braver, 2008). It is a function for selecting and enhancing a limited area of information, while suppressing other areas. Cognitions are included in these areas of information.

To the author's knowledge, connectionist models for cognitive dissonance taking the concepts of self and attention into account have not been presented. In this paper, a novel model considering these concepts is described and reduction of cognitive dissonance based on the model is discussed.

Connectionist Model

Figure 1 shows our connectionist model accounting for cognitive dissonance. In accordance with the cognitive dissonance theory (Festinger, 1957), two cognitions are adopted in the model, which are depicted as units R and I. We assume that unit R is a reality-based cognition such as the cognition of behavior, experience, or actual situation, while unit I is an imagination-based cognition such as the cognition of expectancy, hope, or belief. In the case of the previously mentioned fable, the imagination-based cognition held by the fox is "The grapes can stave off my hunger" and the reality-based cognition is "The grapes do not stave off my hunger."

Attention plays an important role in cognitive systems. It selects and enhances limited cognitions and suppresses others. If we consider two cognitions alone, when one of them is selected and enhanced, the other is rejected and

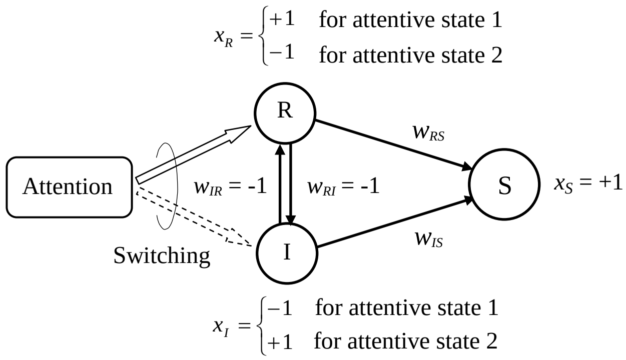


Figure 1: Connectionist model adopting the concepts of attention and self.

suppressed. Thus, the units of these cognitions perform a bistable operation. In order to achieve this operation, the units must be bidirectionally connected through links having negative weights as shown in Figure 1 (Rojas, 1996). In our model, to simplify the discussion, it is assumed that the activation level of units R and I is "+1" or "-1", which correspond respectively to excited and inhibited state, and the weights between the units, i.e., w_{RI} and w_{IR}, are -1. The bistable operation is determined by inputs to the unit pair applied through some additional links connected to the unit pair. The inputs' condition depends on the attentive state. Since attention is usually focused on the cognition of unit R, the bistable operation is not perfectly symmetrical. There might be some bias difference between the inputs to the unit pair. In order to pick up the elements closely related to cognitive dissonance, such additional links are excluded from Figure 1.

Aronson (1969) emphasized that dissonance theory can make the clearest prediction of cognitive dissonance when we deal with the self-concept and cognitions about some behavior. Self is regarded as the key for arousing cognitive dissonance. In accordance with this view, we introduce a unit corresponding to self in our connectionist model, which is depicted as unit S in Figure 1.

Self is a complex system in which cognitive and affective elements are integrated. It is extremely difficult to rigorously represent the self in a simple connectionist model. Here, we assume that self is characterized by static equilibrium, in which there is resistance to incoming information that would change the status of elements (Nowak et al., 2000), and that, as pointed out in many studies of cognitive dissonance, people have a high or positive self-concept. Thus, we boldly introduce a single unit for self whose activation level is constant and takes a value of +1.

Cognition is composed of several elements having different attributes, each of which is related to the self with some evaluation given by the person. Cumulative evaluation of all elements for a cognition is assumed to be the evaluation of the cognition. In our connectionist model, such an evaluation of cognition is represented by the weight of the link between the unit of the cognition and that of self. If the cognition is attractive for self, the evaluation or weight is positive, and if not, negative. The links are assumed to be unidirectional with directions from units R and I to unit S. The weight of the link between units R and S is w_{RS} and that between units I and S is w_{IS}. These weights take values between -1 and +1.

When we respectively represent the activation levels of units R, I, and S by x_R, x_I, and x_S, the model's cognitive dissonance can be described according to the definition given by Shultz and Lepper (1996) as follows:

$$CD = \left(2x_R x_I - w_{RS} x_R x_S - w_{IS} x_I x_S\right)/4 \,. \qquad (1)$$

Focus of attention determines the state of the bistable operation of the unit pair composed of units R and I. We define attentive state 1 as the state in which attention is focused on the cognition of unit R and (x_R, x_I) = (+1, -1), while attentive state 2 is defined as the state in which attention is focused on the cognition of unit I and (x_R, x_I) = (-1, +1). Using Equation (1) and the previously-mentioned assumption that x_S = +1, cognitive dissonances for the two states are described as

$$CD = \begin{cases} (w_{IS} - w_{RS} - 2)/4\,, & \text{for attentive state 1}, \\ (w_{RS} - w_{IS} - 2)/4\,, & \text{for attentive state 2}. \end{cases} \qquad (2)$$

Analysis of Evaluation Change

In neural networks, the weight of a connection between two neurons changes according to the activation condition of the neurons. When the two neurons are excited simultaneously, the weight of the link between them increases and when they are not decreases. We assume that the weights in our connectionist model perform similarly to those of neural networks. The modified Hebbian learning rule presented by Oja (1982) incorporates the saturation characteristics of neurons into the original Hebbian rule. It represents the changes in weight as a function of the activation levels of input and output units, the weight between them, and a constant representing the learning rate during a time interval (O'Reilly & Munakata, 2000).

When we assume that the time interval is infinitesimal, the rule can be represented by the following differential equation:

$$\frac{dw}{dt} = \varepsilon'\left(xy - y^2 w\right)\,, \qquad (3)$$

where ε' is a constant representing the learning rate during a unit time interval.

In order to consider the behavior of w_{RS} we substitute w_{RS} for w in Equation (3). Since $x = x_R = +1$ and $y = x_S = +1$ for attentive state 1 and $x = x_R = -1$ and $y = x_S = +1$ for attentive state 2, we obtain general solutions of Equation(3) as follows:

$$w_{RS} = \begin{cases} K_1 e^{-\varepsilon' t} + 1 & \text{, for attentive state 1,} \\ K_2 e^{-\varepsilon' t} - 1 & \text{, for attentive state 2.} \end{cases} \quad (4)$$

For w_{IS}, similar to the above, we can derive following solutions:

$$w_{IS} = \begin{cases} K_3 e^{-\varepsilon' t} - 1 & \text{, for attentive state 1,} \\ K_4 e^{-\varepsilon' t} + 1 & \text{, for attentive state 2.} \end{cases} \quad (5)$$

K_1, K_2, K_3, and K_4 are constants determined by initial conditions. Since the weights represent evaluations as previously mentioned, Equations (4) and (5) represent the time dependence of evaluations for cognitions R and I.

Comparison with Experimental Results

To confirm the validity of our connectionist model and analysis, here we take up the free-choice paradigm as the first example, and compare theoretical results of our analysis and experimental results reported in the literature. In an experiment carried out by Brehm(1956), subjects were asked to rate each of a variety of items on desirability. They were next required to make a difficult choice, i.e., a choice between two items that they had rated high, or an easy choice, i.e., a choice between one item they had rated high and one they had rated low. The chosen items were given to the subjects who then rated them again. The experimenter measured the differences between the first and second ratings.

Applying our model to the experiment, we regard units R and I as the chosen and rejected items, respectively. This is because after a choice is made, its result is the reality for the subject and the chosen item becomes the element of the reality-based cognition, while the rejected item becomes that of the imagination-based cognition. Weights w_{RS} and w_{IS} are evaluations of the two items. We assume that the choice is carried out at $\varepsilon' t = 0$. Weights' initial values at $\varepsilon' t = 0$ used in our theoretical examination are presented in Figure 2. The theoretical values are determined by making their maximum and minimum possible values, i.e., +1.0 and -1.0, correspond to the experimental maximum and minimum evaluations used in rating, i.e., 8 and 1, respectively. The values in the experiment and those in the theory are linearly related, and their correspondence is schematically shown in Figure 2. Since we consider the situation on the basis of reality, transitions of the weights or evaluations are calculated for attentive state 1.

Figure 3 shows our theoretical results at $\varepsilon' t = 0.7$ as well as the experimental data reported by Brehm(1956). Our theoretical results show that for both the difficult and easy

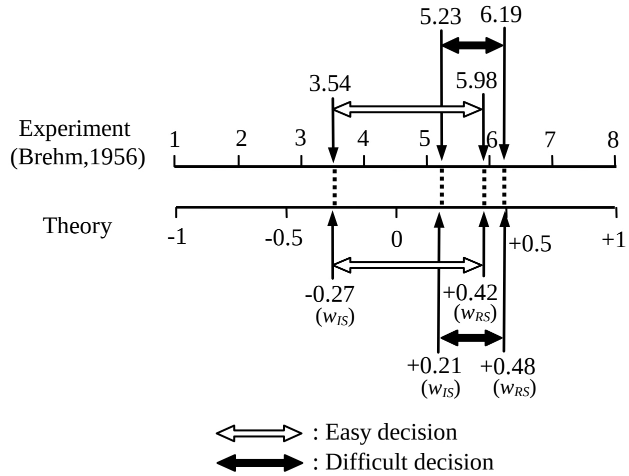
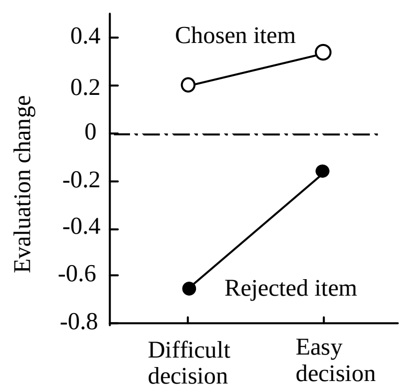
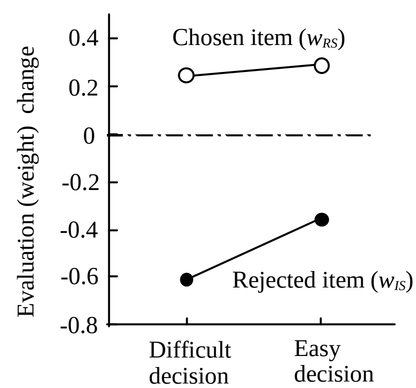


Figure 2 : Correspondence of the initial values of evaluation.



(a) Experiment (Brehm, 1956)



(b) Theory based on our model at $\varepsilon' t = 0.7$

Figure 3: Evaluation change in free-choice paradigm.

choices, w_{RS} increases and w_{IS} decreases, i.e., the separation between them increases, with the increase of time after the choice. The degree of separation increase is substantial for the difficult choice. Our results in Figure 3(b) are similar to the experimental data depicted in Figure 3(a).

We take up the insufficient-justification paradigm as the second example. In an experiment carried out by Freedman (1965), school children were forbidden to play with a desirable toy under either a mild or severe threat and the experimenter either stayed in or left the room while the children played. Actual play with the previously forbidden toy later indicated that derogation was greater under the mild than under the severe threat conditions only when there was no surveillance.

Applying our model to this experiment, we assume that units R and I are respectively the cognitions "I do not play with the toy" and "I play with the toy". This is because, since the children are forbidden to play with the toy and are ordered to obey the directions, their reality-based cognition is "I do not play with the toy." Weights w_{RS} and w_{IS} are respectively evaluations for not playing with and playing with the toy.

Here, we assume that weight w_{IS} is a compound of three subweights $w_0(t)$, w_t, and w_s as shown in Figure 4. The intrinsic evaluation given by the subject for the cognition of unit I is $w_0(t)$ ($-1 \le w_0(t) \le 1$). Since the evaluation may change, it is represented by a function of t. The subweights w_t and w_s are additional weights caused respectively by the effects of threat and surveillance. Since the effects are independent of time, w_t and w_s are constants.

As mentioned above, we assume that the weights are bounded and take values between -1 and +1. Thus, we represent w_{IS} as follows:

$$w_{IS} = \begin{cases} w_0(t) + w_t + w_s, & \text{for } -1 \le w_0(t) + w_t + w_s \le +1, \\ -1, & \text{for } w_0(t) + w_t + w_s < -1, \\ +1, & \text{for } w_0(t) + w_t + w_s > +1. \end{cases} \quad (6)$$

Considering the situation on the basis of reality, we assume the attentive state is 1. Since threat and surveillance are unpleasant for the subject and thus function as negative elements of the cognition represented by unit I, we assume

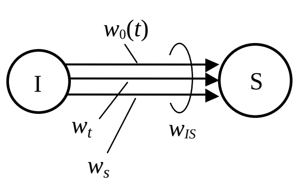


Figure 4: Weights in the connection between units I and S for insufficient-justification paradigm.

Table 1: Weights used in the theory for insufficient-justification paradigm.

	Non-surveillance	Surveillance
Mild threat	w_0 (0) = +1.0 w_t = -0.2 w_s = 0	w_0 (0) = +1.0 w_t = -0.2 w_s = -2.0
Severe threat	w_0 (0) = +1.0 w_t = -1.8 w_s = 0	w_0 (0) = +1.0 w_t = -1.8 w_s = -2.0

$w_t < 0$ and $w_s < 0$. Applying Equation (6) to the theory described in the previous section, we obtain

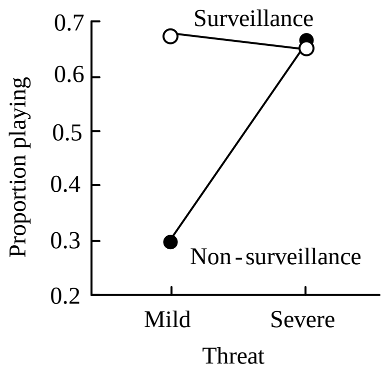
$$w_0(t) = \begin{cases} w_0(0), & \text{for } w_0(0) + w_t + w_s < -1, \\ (w_0(0) + w_t + w_s + 1)\, e^{-\varepsilon' t} - w_t - w_s - 1, \\ & \text{for } w_0(0) + w_t + w_s \ge -1. \end{cases} \quad (7)$$

Table 1 presents the values of the weights (evaluations) used in our theoretical examination. Since the toy might be very attractive for the subjects, we assume that $w_0(0)$ takes the maximum value, i.e., +1.0. As surveillance seems to be effective for forbidding and induce highly negative feeling in the subject, the absolute value of w_s under surveillance is assumed to be large so that w_{IS} takes the minimum value, i.e., -1.0.

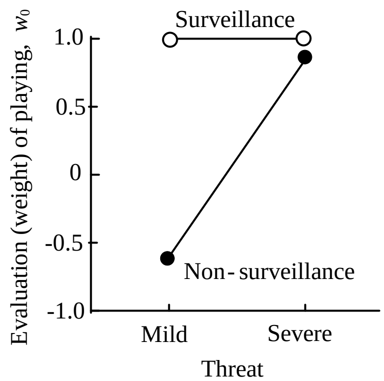
Transitions of the weights or evaluations are calculated for attentive state 1. Figure 5 shows our theoretical results for $w_0(t)$ at $\varepsilon' t = 2$ as well as the experimental data reported by Freedman (1965). The theoretical result shown in Figure 5(b) indicates that derogation is greater under the mild than under the severe threat conditions only when there was no surveillance, which is similar to Freedman's experimental result.

Attention-focus Switching

Here, following the Aesop's fable, we anticipate a case in which a person has a certain expectation concerning something, and he fails to realize it in spite of his effort. The object of expectation is the cognition of unit I in Figure 1, and the result of failure is the cognition of unit R. Since the expected object is attractive for the person and causes positive feeling, the cognition of unit I is positively evaluated and thus we assume w_{IS} >0. In contrast, since the failure result is disagreeable and causes negative feeling, the cognition of unit R is negatively evaluated and thus we assume w_{RS} <0. Since cognitive dissonance occurs when we experience the failure and thus attention is focused on the cognition of unit R, we consider here the situation in



(a) Experiment (Freedman, 1965)



(b) Theory based on our model at $\varepsilon' t$ = 2.0

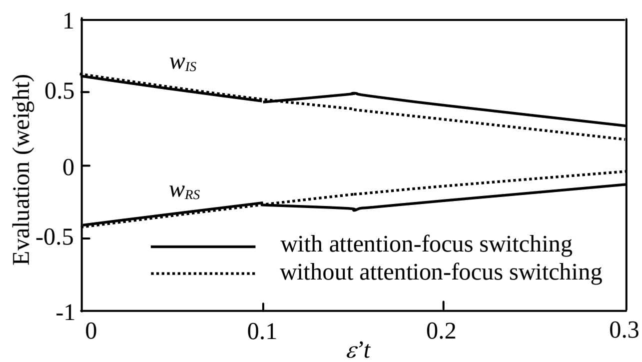
Figure 5: Evaluation change in insufficient-justification paradigm.

attentive state 1. Assuming that, as an example, w_{IS} = +0.6 and w_{RS} = -0.4 are the initial conditions at $\varepsilon' t = 0$, we can obtain w_{IS}, w_{RS}, and cognitive dissonance (*CD*) at $\varepsilon' t \geq 0$ by using Equations (2), (4), and (5), which are shown by the dotted lines in Figure 6. The abscissa is $\varepsilon' t$, which is a parameter proportional to the time passed after the person made the trial and failed.

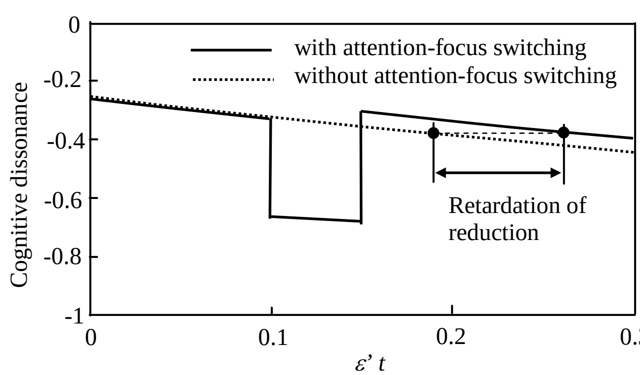
The decrease of w_{IS} and increase of w_{RS} with the increase of time shown in Figure 6(a) indicate adaptation or rationalization of the person under cognitive dissonance. In accordance with these changes, *CD* is reduced with the increase of time as shown in Figure 6(b). Such changes appeared in w_{IS}, w_{RS}, and *CD* corresponds to the phenomenon predicted by the cognitive dissonance theory (Festinger, 1957).

Conventional research on cognitive dissonance does not seem to have amply discussed the effect of attention. Since the occurrence of cognitive dissonance induces unpleasant feeling, the person mentioned above might strive to reduce it. He might recall the expectation and imagine the result that could be obtained in success of his trial. The fox in the fable might say after a short time, “I have gotten hungrier. If I could get the grapes now, even if they were not ripe, I might eat them and be feeling full now”. In this way, the focus of attention is switched from reality-based to imagination-based cognition. However, since one cannot live on imagination only and must act on the basis of reality, it is not long before the focus of attention is switched back to reality-based cognition. Thus, we assume here that the focus of attention is switched and the attentive state is changed such that 1 $\rightarrow$ 2 $\rightarrow$ 1.

Transitions of evaluation and cognitive dissonance under such switching are calculated and depicted by the solid lines in Figure 6. In the first period of $0 \leq \varepsilon' t < 0.1$, where the attentive state is 1, cognitive dissonance is monotonically reduced with the increase of $\varepsilon' t$. At $\varepsilon' t$ = 0.1, the focus of attention is switched to the cognition of unit I and cognitive dissonance is reduced stepwise. During the period of $0.1 \leq \varepsilon' t < 0.15$, where the attentive state is 2, reduction of cognitive dissonance continues with a small reduction rate. At $\varepsilon' t$ = 0.15, the focus of attention is switched back to the cognition of unit R. Cognitive dissonance is then built up stepwise and takes an amount greater than would be taken when switching is not performed at all. After that, attentive state 1 is retained and monotonical reduction of cognitive



(a) Change in evaluation with increase of time.



(b) Change in cognitive dissonance with increase of time.

Figure 6: Changes in evaluation and cognitive dissonance under attention-focus switching.

dissonance continues. Consequently, attention-focus switching causes retardation of cognitive-dissonance reduction as shown in Figure 6(b), which induces lingering of unpleasant feeling or discomfort. If the switching is frequently repeated, cognitive dissonance remains for a long time.

The well-known saying "What's done is done" might imply that it is not worth worrying about an unfavorable situation caused by past behavior. It might also warn of the building-up of cognitive dissonance caused by attention-focus switching and suggest that attention not be focused on imagination-based cognition. Another well-known saying, stated by Dante Alighieri, is that "There is no greater grief than to recall a time of happiness when in misery." This implies that to obtain peace of mind, it is important to focus one's attention on the reality and discard the imagination even if one finds the reality unpleasant. The insistence common to these sayings might support the results shown in Figure 6(b).

Conclusion

This paper described a novel connectionist model accounting for cognitive dissonance, in which the concepts of self as well as attention-focus switching are adopted. The model was investigated not with the computer simulation widely used in conventional research, but with a mathematical analysis based on a differential equation. Predictions based on the model were confirmed to coincide with experimental data reported in the literature. It was shown that attention-focus switching between reality-based and imagination-based cognition causes building-up of cognitive dissonance and retardation of its reduction. This coincides with the implication of well-known sayings suggesting ways to keep the mind away from feelings of suffering or discomfort.

References

Aronson, E. (1969). The theory of cognitive dissonance: A current perspective. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol.4, 1-34. NY: Academic Press.

Brehm, J. W. (1956). Postdecision change in the desirability of alternatives, *Journal of abnormal and social psychology*, *52*, 384-389.

Festinger, L.(1957). *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson.

Freedman, J. L. (1965). Long-term behavioral effects of cognitive dissonance. *Journal of Experimental Social Psychology*, *1*, 145-155.

Mosler, H., Schwartz, K., Ammann, F., & Gutsher, H. (2001). Computer simulation as a method of further developing a theory: Simulating the elaboration likelihood model, *Personality and Social Psychology Review*, *5*(3)201-215.

Nowak, A., Vallacher, R. R., Tesser, A., & Borkowski W. (2000). Society of self: The emergence of collective properties in self-structure. *Psychological Review*, *107*, 39-61.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, *15*, 267-273.

O'Reilly, R. C., & Munakata, Y. (2000). Computational explorations in cognitive neuroscience. Cambridge, MA: MIT Press.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion*. NY: Springer-Verlag.

Pisapia, N. D., Repovs, G., & Braver, T. S. (2008). Computational models of attention and cognitive control. In R. Sun (Ed.), *The Cambridge handbook of computational psychology*, NY: Cambridge University Press.

Read, S. J., & Miller, L.C. (1994). Dissonance and balance in belief systems: The promise of parallel constraint satisfaction processes and connectionist modeling approaches. In R. C. Schank & E. Langer (eds.), *Beliefs, reasoning, and decision-making: Psycho-logic in honor of Bob Abelson*. Hillside, NJ: Lawrence Erlbaum.

Rojas, R. (1996). Neural networks: A systematic introduction, Berlin: Springer-Verlag.

Shultz, T. R., & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review*, *103*, 219-240.

Shultz, T. R., Leveille, E. & Lepper, M. R. (1999). Free choice and cognitive dissonance revisited: Choosing "Lesser Evils" versus "Greater Goods". *Personality and Social Psychology Bulletin*, *25*, 40-48.

Spellman, B. A., Ullman, J. B., & Holyoak, K. J. (1993). A coherence model of cognitive consistency: Dynamics of attitude change during the Persian Gulf War. *Journal of Social Issues*, *49*(4), 147-165.

Thibodeau, R. & Aronson, E. (1992). Taking a closer look: Reasserting the role of the self-concept in dissonance theory. *Personality and Social Psychology Bulletin*, *18*, 591-602.

Van Overwalle, F., & Jordan, K. (2002). An adaptive connectionist model of cognitive dissonance. *Personality and Social Psychology Review*, *6*, 204-231.

Van Overwalle, F., & Siebler, F. (2005). A connectionist model of attitude formation and change. *Personality and Social Psychology Review*, *9*, 231-274.