Consonance Network Simulations of Arousal Phenomena in Cognitive Dissonance

Thomas R. Shultz (shultz@psych.mcgill.ca)
Department of Psychology; McGill University
Montreal, QC H3A1B1 Canada

Mark R. Lepper (lepper@psych.stanford.edu)
Department of Psychology; Stanford University
Stanford, CA 94305-2130 USA

Abstract

The consonance constraint satisfaction model, recently used to simulate the major paradigms of cognitive dissonance theory, is extended to deal with emotional arousal phenomena in dissonance. The impact of arousing drugs is implemented in the simulations by a scalar that modulates the intensity of unit activations representing the relevant cognitions and the connection weights representing their implications. The simulations show that even exotic dissonance phenomena can be explained in terms of the relatively common process of constraint satisfaction.

For centuries, cognitive consistency has been a battleground in the continuing philosophical debates over human rationality. Consistency was often seen as a hallmark of reason, or as a necessary condition for a set of beliefs or arguments to be considered logical. Alternatively, consistency was dismissed as wooden-headed conservatism—the "hobgoblin of small minds" or the "last refuge of the unimaginative."

Within psychology, cognitive consistency has similarly played a major role in debates about human reason (Abelson, 1971; Lepper, 1994). In many of the early consistency models, pressures toward balance (Heider, 1946, 1958), congruity (Osgood & Tannenbaum, 1955), or symmetry (Newcomb, 1953) in a person's cognitive system were viewed as expressions of a logical need to achieve a coherent understanding of a sometimes contradictory world (McGuire, 1960). In contrast, other consistency models, especially Festinger's (1957) theory of cognitive dissonance, focused on the ways in which the human desire to avoid inconsistency (dissonance) could lead to behavior that appeared fairly irrational. In this latter tradition, people were seen more as "rationalizing," than as "rational," creatures (Aronson, 1969). This emphasis on irrationality in dissonance theory contributed to the view that dissonance phenomena were noticeably different, even somewhat exotic, compared to other everyday psychological phenomena.

After a considerable period of relative quiescence, interest in cognitive consistency in general, and cognitive dissonance in particular, has recently been rekindled, from two quite disparate directions (e.g., Harmon-Jones & Mills, 1999).

One source of renewed interest in cognitive consistency has come from attempts to model relationships among social attitudes and beliefs in computational terms (e.g., Read & Miller, 1994, 1998; Shultz & Lepper, 1996, 1998; Spellman, Ullman, & Holyoak, 1993). Shultz and Lepper (1996, 1998), for example, modeled a variety of phenomena from some of the central and most robust cognitive dissonance paradigms. Their so-called "consonance" model uses connectionist neural networks operating by the principle of constraint satisfaction to simulate traditional findings in cognitive dissonance. Their model captured the results of a number of classic "insufficient justification" phenomena, including the effects of threats of punishment on liking for a forbidden toy (e.g., Freedman, 1965), the consequences of "forced compliance" with a request to engage in counter-attitudinal activities (e.g., Linder, Cooper, & Jones, 1967), and the psychological effects of initiations (e.g., Gerard & Mathewson, 1966). It also captured basic phenomena concerning the consequences of making a free choice (e.g., Brehm, 1956) and predicted new free-choice effects that were subsequently confirmed in further psychological experimentation (Shultz, Léveillé, & Lepper, 1999).

These simulations suggested that even the apparently exotic, counter-intuitive, and irrational effects emphasized by early dissonance theorists could be interpreted in terms of considerably more mundane cognitive processes. In several cases, psychological phenomena were covered more accurately in the simulations than they were by classical dissonance theory. The superior coverage of the consonance model was due to the inclusion of constraints not present in dissonance theory and to the increased precision inherent to a computational formulation. The success of the consonance model allows a reinterpretation of cognitive dissonance and its reduction that emphasizes what it has in common with many other psychological phenomena operating according to constraint satisfaction principles (Holyoak & Thagard, 1989; Kintsch, 1988; Read & Miller, 1994, 1998; Rumelhart, Smolensky, McClelland, & Hinton, 1986; Sloman, 1990; Spellman & Holyoak, 1992; Spellman et al., 1993; Thagard, 1989).

A second main source of the recent resurgence of interest in consistency phenomena has come from social psychologists wanting to highlight and investigate the affective and motivational properties of cognitive dissonance (Elliot & Devine, 1994; Harmon-Jones & Mills, 1999). Some of these efforts involved attempts to study the critical role of actual physiological arousal in the production and reduction of dissonance (e.g., Cooper & Fazio, 1984).

The present paper attempts to integrate these two newer directions in cognitive dissonance research by presenting simulations, based on the consonance model, of key phenomena in the more recent dissonance literature concerning the emotionally-arousing properties of cognitive dissonance. Simulations of this sort can provide a relatively precise and reliable extension of theory to explain empirical psychological findings (Smith, 1996).

Arousal and Dissonance

Even in the earliest dissonance research, it was postulated that dissonance involves a state of aversive affective arousal (Festinger, 1957). In more recent years, a number of experiments supported this notion (Cooper & Fazio, 1984). Thus, dissonance energizes dominant responses, just as other arousal states do (Pallak & Pittman, 1972), and produces actual physiological arousal (Croyle & Cooper, 1983). Even more dramatically, the presence and size of dissonance effects vary with independent manipulations of arousal.

For example, Cooper, Zanna, and Taves (1978) directly manipulated arousal with drugs and showed the necessity of arousal for subsequent attitude change. Cooper et al.'s participants were asked to write an essay that went against their own attitudes. The essay favored pardoning former President Richard Nixon—a policy with which virtually all participants strongly disagreed. The participants were given either high or low choice to write these counter-attitudinal essays, under three different drug conditions. In the context of an earlier experiment, however, the participants had each just taken a pill that they had been told was a harmless placebo. Actually, however, the drug was either a placebo, a "downer" (Phenobarbital), or an "upper" (amphetamine).

Results from the Cooper et al. study are shown in Figure 1. In the placebo condition, there was the usual dissonance effect, i.e., more attitude change in the direction of the essay under high choice than under low choice. This effect of choice was eliminated in the "downer" condition, where there was very little attitude change regardless of choice, and enhanced in the "upper" condition, producing a significant dissonance effect even under low choice. These results became the focus of our simulations.

The Consonance Model

The consonance model is based on the idea that dissonance reduction can be interpreted as a constraint satisfaction problem. The motivation to seek cognitive consistency that is postulated by dissonance theory and related theories can be viewed as imposing constraints on the beliefs and attitudes that an individual holds at any given moment (e.g., Abelson, Aronson, McGuire, Newcomb, Rosenberg, & Tannenbaum, 1968; Abelson & Rosenberg, 1958; Feldman, 1966). Such consistency problems can be solved by satisfying multiple soft constraints, conditions that are desirable, but not essential, to satisfy and which may vary in their relative importance to the individual.

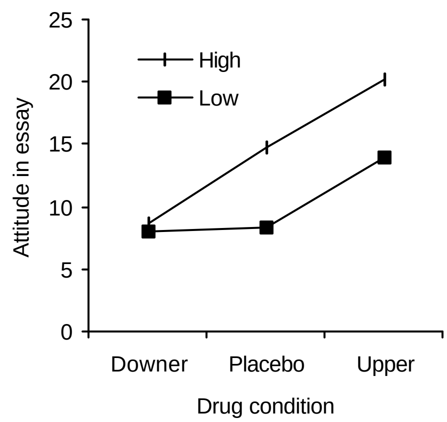

Figure 1: Mean attitude as a function of choice and drug (from Cooper et al., 1978).

In this model, consonance networks correspond to an individual's representation of the situation created in the various conditions of a cognitive dissonance experiment. Unit activations represent the direction and strength of the person's attitudes and beliefs. Units can also differ in their resistance to activation change, reflecting differences in the extent to which particular cognitions are supported by other cognitions or are clearly anchored in reality. Connection weights between different cognitions represent psychological implications among the person's attitudes and beliefs. The connections between any two units can be excitatory, inhibitory, or nonexistent. Both unit activations and connection weights can vary across the different conditions of a single experiment.

Consonance is roughly the degree to which similarly evaluated units are linked by excitatory weights and oppositely valued units are linked by inhibitory weights. More formally, the consonance contributed by a particular unit i is:

$$consonance_i = \sum_j w_{ij} a_i a_j \tag{1}$$

where w_{ij} is the weight between units i and j, a_i is the activation of the receiving unit i, and a_j is the activation of the sending unit j. The overall consonance in a given network is the sum of the values given by Equation 1 over all receiving units in that network:

$$consonance_n = \sum_i \sum_j w_{ij} a_i a_j \tag{2}$$

Activations change over time cycles in order to satisfy constraints and increase consonance. Activation spreads over time cycles by two update rules:

$$a_i(t+1) = a_i(t) + net_i(ceiling - a_i(t)) \text{ , when } net_i \geq 0 \quad (3)$$

$$a_i(t+1) = a_i(t) + net_i(a_i(t) - floor) \text{, when } net_i < 0 \quad (4)$$

where $a_i(t+1)$ is the activation of unit i at time $t + 1$, $a_i(t)$ is the activation of unit i at time t, *ceiling* is the maximal level of activation, *floor* is the minimal activation, and net_i is the net input to unit i, defined as:

$$net_i = resist_i \sum_j w_{ij} a_j \quad (5)$$

The parameter $resist_i$ indexes the resistance of receiving unit i to having its activation changed.

At each time cycle during the simulation, n units are randomly selected and updated according to Equations 3 and 4. For most of our simulations, n is the number of units in the network. The update rules described in Equations 3-5 ensure that consonance increases or stays the same across time cycles. When consonance reaches an asymptote, the updating process is stopped.

The design of consonance networks to implement particular dissonance experiments follows a set of five principles that map cognitive dissonance theory to the consonance model. Principle 1 specifies that a cognition is implemented by the net activation of a pair of negatively connected units, one representing the positive pole and the other representing the negative pole. This permits the model to deal with both conflict and ambivalence. Net activation for the cognition is the difference between activation of the positive unit and activation of the negative unit. Activations range from a floor to a ceiling. In our dissonance simulations, the floor parameter has a default value of 0. The ceiling parameter is 1 for positive poles, and 0.5 for negative poles.[1]

Principle 2 specifies that connections between cognitions are based on inferred causal implications between those cognitions (Abelson, 1968). Connection weights range from -1 to +1, with 0 representing a lack of causal relation. When two cognitions are positively related, their positive poles are connected with excitatory weights, as are their negative poles. Inhibitory weights connect the positive pole of one cognition with the negative pole of the other cognition. These connections are exactly reversed for cognitions that are negatively related. Each unit has an inhibitory self-connection specified by the *cap* parameter, and all connection weights are bi-directional. Connection weights have a default value of 0.5.

Principle 3 specifies that the total amount of dissonance in a network is the negative of total consonance divided by r, which is the number of nonzero relations among cognitions:

[1] When neurons are organized into excitatory and inhibitory groups that respond in opposite ways to the same input, the activation range for excitatory neurons is usually greater for positive than for inhibitory neurons (Anderson, 1995, pp. 150-152).

$$dissonance = \frac{-\sum_i \sum_j w_{ij} a_i a_j}{r} \quad (6)$$

Dividing by r standardizes dissonance across networks by controlling for the number of relevant relations. Self-connections are excluded from this computation of dissonance. This definition of dissonance differs from Festinger's (1957) definition, not only because it is formalized, but also because it measures the amount of dissonance within each inter-cognition relation, includes within-cognition ambivalence, and varies even when all relations are dissonant or all relations are consonant.

Principle 4 maps dissonance reduction to activation updates by specifying that networks settle into more stable, less dissonant states as unit activations are updated with Equations 3, 4, and 5. The *cap* parameter with a default value of -0.5, corresponding to the value of the connection between each unit and itself (w_{ii}), prevents activations from reaching the ceiling.

Principle 5 specifies that cognition unit activations, but not connection weights, are allowed to change, and that some cognitions are more resistant to change than others, as implemented in Equation 5. Typically, beliefs, behaviors, and justifications are more resistant to change than are evaluations or attitudes. The *resist* parameter has default values of 0.5 for low resistance and 0.01 for high resistance. As specified in Equation 5, the larger the resistance multiplier, the more readily the unit changes its activation.

In order to assess the robustness of simulation results across variability in the specific parameters, weights, resistances, caps, and initial activations are all randomized by adding or subtracting a random proportion of their initial amounts. The *rand%* parameter specifies the proportion in which additions or subtractions are randomly selected with a uniform distribution. Randomization increases psychological realism in the sense that not everyone can be expected to have precisely the same parameter values. Randomization of weight values violates connection weight symmetry, such that $w_{ij} \neq w_{ji}$, and increases instability of network solutions (Hopfield, 1982, 1984). We often compare low (0.1), medium (0.5), and high (1.0) levels of *rand%* in order to assess the stability of network solutions (Shultz & Lepper, 1996). All of the default values mentioned above are consistent across all of our simulations. Additional details about the consonance model and discussions of its assumptions are available in other sources (Shultz & Lepper, 1996, 1998).

The Present Simulations

Network Design

The present simulations focused on the Cooper et al. (1978) experiment described earlier. Because there were no differential payments, as is typical in forced compliance experiments (e.g., Collins, 1973; Linder et al., 1967), there are only two cognitions to model: the attitude and the counter-attitudinal essay. Initially, attitude is given a high negative activation (-0.5) because it is strongly against pardoning Nixon. The essay cognition starts with a high positive activation (0.5) because it is in favor of pardoning Nixon. Using Principle 2, we implement degree of choice by varying the strength of the connection between attitude and essay: high (0.5) vs. low (0.1). Both relations are positive because the more favorable one's attitude, the more likely it would be for one to support this position in writing.[2] Following Principle 5, writing the essay is given high resistance to change because it is a public and irrevocable behavior; attitude is given low resistance to change because it is a subjective evaluation.

The various drug conditions are implemented with a scalar value that multiplies the initial values of connection weights and cognitions: 1.0 for the placebo, 0.5 for the "downer," and 1.5 for the "upper" condition. The basic idea is that these arousal-modulating drugs dampen or boost everything in the system, connection weights and initial activations. This interpretation is consistent with the energizing properties of dissonance (Cooper & Fazio, 1984). The fact that even the low choice/"upper" participants said in manipulation checks that they felt a high level of choice (Cooper et al., 1978) underscores the importance of scaling the connection weights. This scaling is also consistent with evidence that phenobarbital depresses, whereas amphetamine increases, neural firing rates and synaptic transmission (Quastel, 1975). Initial network values for the placebo condition obviously do not change with a multiplier of 1.0.

Twenty networks were run in each of the six experimental conditions, each network having somewhat different parameter settings and a different randomly determined pattern of activation updates. Each network ran for 30 cycles, because dissonance and unit activation values typically reached asymptote by that time.

Results

Mean attitudes towards the view espoused in the essay are presented in Figure 2 at the lowest level of parameter randomization (0.1). The drug x choice interaction resembles that produced in Cooper et al.'s (1978) study. We performed an ANOVA in which drug, with three levels, and choice, with two levels, served as factors. As predicted, there was an interaction between drug and choice, $F(2, 114) > 900$, $p < .001$. The precise nature of this interaction was assessed using a contrast F with weights suggested by Cooper et al.'s human data. The contrast weights were -1, 0, and +3 for the high choice "downer," placebo, and "upper" conditions, respectively; and -1, -1, and 0 for the low choice "downer," placebo, and "upper" conditions, respectively. The contrast was highly significant $F(1, 114) > 12000$, $p < .001$, accounting for 89% of the between-conditions variance.

To test for a significant dissonance effect, even under low choice, when amphetamines had been administered, we computed a regression F for only the three low choice cells, using weights of 0 for the high choice cells, and weights of -1, -1, and 2 for the low choice "downer," placebo, and "upper" conditions, respectively, $F(1, 114) > 500$, $p < .001$. Just as with Cooper et al.'s participants, a dissonance effect was found in the low choice, "upper" condition.

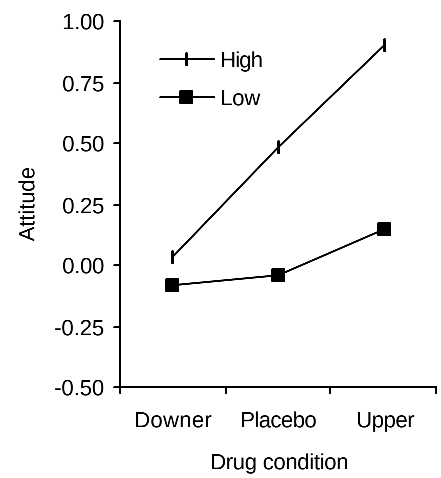

Figure 2: Mean simulated attitude as a function of choice and drug condition.

Mean simulated dissonance scores for the six conditions of the Cooper et al. experiment are plotted in Figure 3 over time cycles. A substantial amount of dissonance in the high choice "upper" condition is strongly reduced, a moderate amount of dissonance in the high choice placebo condition is moderately reduced, and a small amount of dissonance in the low choice "upper" condition is slightly reduced. The other three conditions show almost no dissonance and almost no dissonance reduction. All of these results held up at medium (0.5) and even at high (1.0) levels of parameter randomization.

Discussion

The consonance network simulations fit the human data from Cooper et al. quite precisely. There was the typical dissonance effect in the placebo condition (i.e., more attitude change in the direction of the views expressed in the essay under high choice than under low choice), very little attitude change in the "downer" condition, and enhanced attitude change in the "upper" condition. The only required change from our previous forced-compliance simulations was the inclusion of a scalar parameter to multiply initial values of activations and weights. Consistent with psychological and neurophysiological evidence, this scalar enhanced activations and weights in the "upper" condition

[2] In our previous simulation of the Linder et al. (1967) forced compliance experiment, which had no choice rather than low choice, we cut this link between attitude and essay to zero (Shultz & Lepper, 1996).

and dampened them in the "downer" condition, relative to the placebo control condition. Plots of dissonance reduction suggest that the amount of attitude change is a direct function of the amount of dissonance reduced in the networks.

It appears that as though the basic phenomena on arousal in dissonance can be captured with constraint satisfaction consonance networks, again suggesting that dissonance arousal and reduction has much in common with other constraint satisfaction processes.

Acknowledgments

This research was supported by a grant to the first author from the Social Sciences and Humanities Research Council of Canada and by grant MH-44321 to the second author from the U.S. National Institute of Mental Health. We are grateful for comments on an earlier draft by Sylvain Sirois and Dave Buckingham.

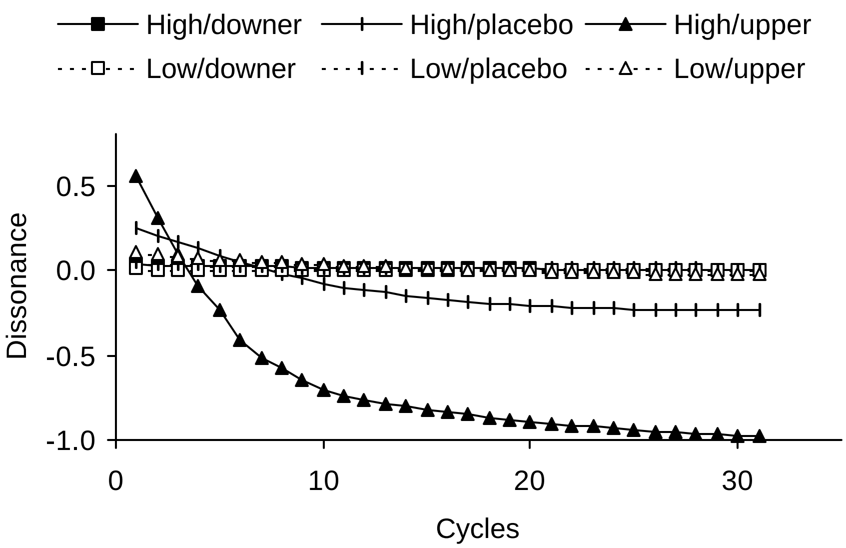

Figure 3: Simulated dissonance over time cycles.

References

Abelson, R. P. (1968). Psychological implication. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook*. Chicago: Rand McNally.

Abelson, R. P. (1971). Are attitudes necessary? In B. T. King & E. McGinnies (Eds.), *Attitudes, conflict, and social change*. New York: American Marketing Association.

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. H. (Eds.) (1968). *Theories of cognitive consistency: A sourcebook*. Chicago: Rand McNally.

Abelson, R. P., & Rosenberg, M. J. (1958). Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science, 3*, 1-13.

Anderson, J. A. (1995). *An introduction to neural networks.* Cambridge, MA: MIT Press.

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

Brehm, J. W. (1956). Post-decision changes in the desirability of choice alternatives. *Journal of Abnormal and Social Psychology, 52*, 384-389.

Collins, B. E. (1973). *Public and private conformity: Competing explanations by improvisation, cognitive dissonance, and attribution theories*. Andover, MA: Warner Modular Publications.

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17). New York: Academic Press.

Cooper, J., Zanna, M. P., & Taves, P. A. (1978). Arousal as a necessary condition for attitude change following forced compliance. *Journal of Personality and Social Psychology, 36*, 1101-1106.

Croyle, R., & Cooper, J. (1983). Dissonance arousal: Physiological evidence. *Journal of Personality and Social Psychology, 45*, 782-791.

Elliot, A. J., & Devine, P. G (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology, 67*, 382-394.

Feldman, S. (Ed.) (1966). *Cognitive consistency*. New York: Academic Press.

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.

Gerard, H. B., & Mathewson, G. C. (1966). The effects of severity of initiation on liking for a group: A replication. *Journal of Experimental Social Psychology, 2*, 278-287.

Harmon-Jones, E., & Mills, J. (Eds.). (1999). *Cognitive dissonance: Progress on a pivotal theory in social psychology*. Washington, DC: American Psychological Association.

Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology, 21*, 107-112.

Heider, F. (1958). The psychology of interpersonal relations. New York: Wiley.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295-355.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA, 79*, 2554-2558.

Hopfield, J. J. (1984). Neurons with graded responses have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences, USA, 81*, 3088-3092.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163-182.

Lepper, M. R. (1994). “Hot” vs. “cold” cognition: An Abelsonian journey. In R. Schank & E. Langer (Eds.), *Beliefs, reasoning, and decision making: Psycho-logic in honor of Bob Abelson*. Hillsdale, NJ: Erlbaum.

Linder, D. E., Cooper, J., & Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology, 6*, 245-254.

McGuire, W. J. (1960). A syllogistic analysis of cognitive relationships. In W. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm (Eds.), *Attitude organization and change*. New Haven, CT: Yale University Press.

Newcomb, T. M. (1953). An approach to the study of communicative acts. *Psychological Review, 60*, 393-404.

Osgood, C. E. & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review, 62*, 42-55.

Pallak, M. S., & Pittman, T. S. (1972). General motivation effects of dissonance arousal. *Journal of Personality and Social Psychology, 21*, 349-358.

Quastel, J. H. (1975). Effects of drugs on energy metabolism of the brain and on cerebral transport. In L. L. Iverson, S. D. Iverson, & S. H. Snyder (Eds.), *Handbook of psychopharmacology: Vol. 5. Synaptic modulators*. New York: Plenum.

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*. Hillsdale, NJ: Erlbaum.

Read, S. J., & Miller, L. C. (Eds.). (1998). *Connectionist models of social reasoning and social behavior.* Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. (1986). Schemata and sequential thought processes in PDP models. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2). Cambridge, MA: MIT Press.

Shultz, T. R., & Lepper, M. R. (1996). Constraint satisfaction modeling of cognitive dissonance phenomena. *Psychological Review, 103*, 219-240.

Shultz, T. R., & Lepper, M. R. (1998). The consonance model of dissonance reduction. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior.* Hillsdale, NJ: Erlbaum.

Shultz, T. R., Léveillé, E., & Lepper, M. R. (1999). Free-choice and cognitive dissonance revisited: Choosing “lesser evils” vs. “greater goods.” *Personality and Social Psychology Bulletin*, *25*, 40-48.

Sloman, S. (1990). *Persistence in memory and judgment: Part-set inhibition and primacy*. Unpublished Doctoral Dissertation, Stanford University.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology, 70*, 893-912.

Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler, then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology, 62*, 913-933.

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*, 147-165.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12*, 435-502.