A Computational Model of Reactive Depression

Charles Webster[1,2], MD, Richard M Glass, MD[3], Gordon Banks, PhD, MD[1]

[1]Decision Systems Lab, Dept. of Neurology, U of Pittsburgh
[2]Lab for Computational Linguistics, Dept. of Philosophy, CMU
[3]Dept. of Psychiatry, U of Chicago

Introduction

We propose a learning model of normal reactive depression and describe DEP (Depression Emulation Program), a computer simulation of selected aspects of information processing in depression. Behaviors we address include the role of stable, internal, and global failures in triggering depression, increased objectivity, the cognitive loop, fluctuations of self-describing adjectives, and decreased motivation. These behaviors are consistent with the hypothesis that reactive depression involves controlled processing of the self-schema in order to learn new automatic emotional processes for directing attention to goals and plans.

Occasional mild depression may be an adaptive response to a changing environment (Klerman, 1987). Emotions narrow attention to personally relevant inferences (de Sousa, 1987). In natural language understanding, emotional context provides powerful constraints for understanding the goals and plans of characters in narrative (Dyer, 1983). In cognitive psychology, emotions have been hypothesized to motivate the selection of situationally and personally relevant goals (Simon, 1967). We theorize that the nondepressed person goes about his daily activities with largely automatic emotional responses directing attention to goals and plans. A depressed person, however, seems to require much effort in order to carry out the same activities. After a failure (losing a job, for example), a person may have to adopt new automatic strategies, and reactive depression may actually facilitate the discarding of obsolete behaviors and the formation of new ones. During this reformation, automatic processing of the self-schema may be suspended in favor of more slow, effortful, and sequential processing.

DEP is a program embodying a computational learning theory of adaptive reactive depression (transient depressed mood following an environmental event; all references to depression herein will be to this specific type). In the nondepressed state, DEP directs attention rapidly and effortlessly (automatic processing), without environmental feedback. In the depressed state, DEP directs attention slowly and effortfully (controlled processing), while environmental feedback modifies attention-directing emotional responses. DEP models attention direction using a connectionist discrimination network, DEP's analog of the self-schema. Automatic network descent is achieved when each node in the descent path has sufficient activation. In controlled descent, insufficient activation leads to slow sequential search for a correct descent path. When a correct path is found, DEP's weights are adjusted to increase the activation level of the path in response to future inputs. Thus, controlled descent of DEP's self-schema is instrumental in the development of future correct automatic responses. DEP is not based on a theory of clinical depression. However, a model of "normal", reactive depression may be a useful intermediate step toward a computational model of "pathological" depression.

Automatic and Controlled Emotional Processing of the Self-Schema

We use two concepts from cognitive psychology in our model of depressive information processing: the self-schema (Marcus, 1980), and automatic and controlled processing (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Schneider et al., 1984; Schneider & Detweiler, 1987). The self-schema is a hierarchical representational system for encoding, organizing, and retrieving declarative and procedural information about the self. It goes under a variety of names: self-concept (Shavelson, 1986), system concept (Carver & Scheier, 1986), self-schemata (Marcus, 1980), and self-schema (Coyne & Gotlib, 1983). We hypothesize that the self-schema can exist in an automatic (nondepressed) or controlled (depressed) state. An automatic process occurs as a "fast, parallel, effortless process that is not under the direct subject control, and is responsible for the performance of well-developed skilled behavior." (Schneider et al., p. 1, 1984). Controlled processing is a "slow, generally serial, effortful, capacity-limited, subject-regulated processing mode that must be used to deal with novel or inconsistent information." (Schneider et al., p. 2, 1984). Learning new automatic processes requires controlled processing.

We conceive of the self-schema as a hierarchy of attributes and procedures enabling a person to select goals and plans specialized for important social roles (act attentive in class, take notes, act cultured at the opera, comment on the quality of the singing, etc.) based on stable self-attributes (degrees of intelligence, attractiveness, lovableness, etc.). "Motivated" (intensely and persistently goal-directed) behavior occurs when an individual automatically traverses the self-schema from the top to the bottom to arrive at a leaf node with executable plans. Figure 1 is a prototypical self-schema. Environmental features such as it being a workday and being hungry activate nodes in the self-schema. The weights between the upper nodes and the intermediate nodes encode self-attributes, such as "hardworking" and "ambitious". The terminal nodes activate plans like getting out of bed, going to work, and asking the boss to lunch. The self-schema is used like a discrimination network (Feigenbaum & Simon, 1984). For example, a person whose self-schema strongly encodes "hardworking" and "ambitious" self-descriptions will automatically go to work and ask the boss to lunch if it is a weekday lunchtime.

Deciding which arc to take during downward traversal of the self-schema requires an "emotional" reaction to plan outcomes (would the outcome make one happy, unhappy, frustrated, angry, etc.?). The path selected depends on outcomes of previously selected goals and plans in similar situations. Normally the traversal from the top to the bottom of the self-schema is automatic and self-schema behaves in a stereotypical manner, as in familiar social roles, effortlessly directing attention to self-relevant goals and plans, and creating default expectations for the results of actions. However, if the individual fails in an important social role, perhaps habitual behavior was not flexible enough, attention was misdirected, or wrong expectations were generated. Personal failure (for example, asking the boss to lunch and being told you are in danger of being fired) is normally a novel situation requiring a flexible response involving controlled processing. Counter-acting the failure may require controlled processing to block, change, or trigger automatic processes. Creation of new, less failure prone, automatic emotional processes requires controlled processing of the self-schema.

A major theory of depressive cognition can be understood in this light. According to the revised learned-helplessness model (Seligman, 1984), depression occurs when people explain negative personal outcomes in terms of *stable*, *internal*, and *global* factors. Failure usually requires specific attributions to trigger depression. In terms of our hierarchical self-schema and its role in focusing attention on appropriate goals and plans such a mechanism makes sense. We would only want to change such a representation if the failure is attributed to stable factors ("My performance will continue to disappoint the boss"), internal factors ("I could have prevented his dissatisfaction"), and to global factors ("I have similar problems in other areas

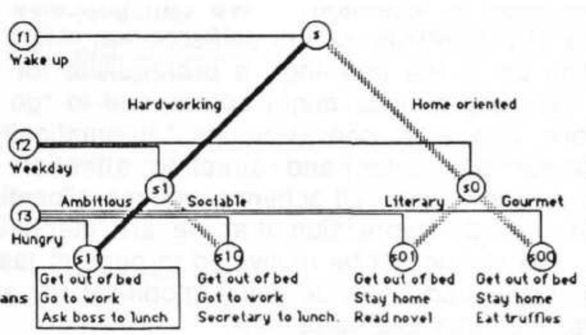

Figure 1: Prototypical Self-Schema

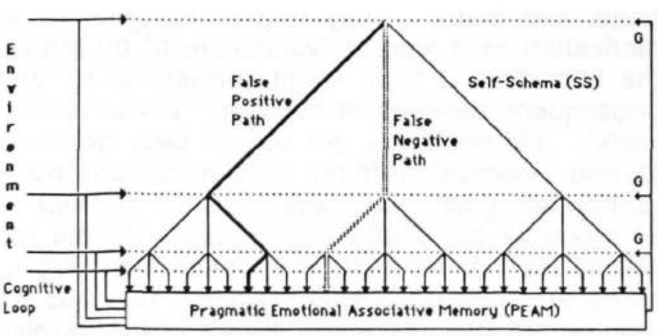

Figure 2: DEP Architecture

of my life"). To the degree that failure meets these criteria, controlled processing should be allowed to modify and reautomate how we see ourselves and what policies we pursue.

A cognitive model of depression should explain the *cognitive loop*, the sequence of failure-related memories and thoughts consuming the depressed person's attention. This cognitive loop has been explained in terms of the network structure of episodic and emotional memory (Bower, 1981; Ingram, 1984), and the evoking of emotionally related memories (Blaney, 1986). Memories acquired during a particular emotional state become attached to emotion "nodes". Experiencing an emotion activates memories associated with the emotion, leading to "mood congruent" memory retrieval. A failure produces the subjective experience of being reminded of a series of unpleasant memories involving similar failures. Consciously experiencing each unpleasant memory evokes another mood congruent memory and the cycle becomes self-perpetuating.

What is the significance of the cognitive loop? One answer may lie in the role of repetition in automation. While controlled processes may suggest changes in the self-schema, for these changes to become automatic the self-schema must be trained against a database of relevant examples. Given a self-schema whose structure is based on pragmatic and emotional knowledge, mood congruent memory retrieval may serve as an index into memories of previous events most similar to the personal failure triggering the depression. The cognitive loop occurs because of the repetitive nature of "tuning" the self-schema against a database of test cases.

Support for the view that reactive depression involves a malleable self-schema can be found in the phenomenon of *slot rattling*. When asked to assess themselves in terms of a list of adjectives, the mildly depressed individual selects more negative adjectives on average, however day-to-day fluctuations occur (Ross, 1985). We would reasonably expect this variation if these adjectives reflected changing generalizations about ourselves.

A model of depression as the inductive restructuring of the self-schema partially explains an interesting observation. According to a major theory of depressive cognition, negative schemata distort experience to support a negative view of the depressed individual, the situation, and the future (Beck, 1979). While this theory is formulated with respect to clinical depression, it is intriguing that mildly depressed people are actually more *objective* about their performance at a variety of tasks than nondepressed people (Alloy & Abramson, 1979; Coyne & Gotlib, 1983; Giles & Shaw, 1987). This may be explained by hypothesizing that mildly depressed people pay more attention to information relevant to the modification and reautomation of the self-schema. Objective assessments about how well a task is performed are part of the multitude of observations used to inductively reconstruct the self-schema. Metaphorically, the self-schema is a hierarchical edifice resting on a foundation of observations of individual interactions with the world. Like the construction of a scientific theory, its validity rests on the objectivity of individual experimental observations.

A ubiquitous phenomena in depression is *loss of motivation*. In our model, this loss of motivation corresponds to the loss of automaticity of the descent of the self-schema. Automatic emotional responses to a situation cease to automatically and effortlessly formulate appropriate goals and plans. They require the effortful allocation of attention. We can conceive of motivation as a kind of "volunteerism" on the part of subportions of the self-schema. High in the hierarchy, a node might correspond to "getting up in the morning", a prerequisite for all subsequent activities of the day. Lower in the hierarchy a node might correspond to "go to work". Normally we get out of bed, go to work, and start conversations "automatically". During depression these responses are no longer automatic, and directing attention to appropriate goals and plans becomes *effortful*. Descending the self-schema requires allocation of attention away from other tasks. In the mild periodic depression that we are theorizing about, this may be more of a feature than a flaw. We should not be motivated to perform tasks using an attention-directing system that has recently failed, is undergoing modification, and may be preoccupied with a stream of failure-related thoughts and memories.

Depression Emulation Program

DEP is a computer simulation consistent with two sets of constraints. It must reflect the central hypothesis that reactive depression is a form of failure-triggered controlled processing of the self-schema. It must also exhibit a set of empirically observed depressive behaviors. We will describe DEP's architecture while illustrating its gross behavior.

DEP has two subprograms: SS, for Self-Schema, and PEAM, for Pragmatic-Emotional Associative Memory. SS is a connectionist discrimination network with environmental inputs (Figure 2, SS). The weights connecting SS's nodes with each other and the environment represent the potentials for emotional reactions elicited by inputs from the environment. A descending, activated path to a leaf node is analogous to an emotional state (for example, the dark path in Figure 2), and the associated goals, plans, and expectations are accessed by activating the leaf node. Descent is either "automatic" or "controlled", depending on whether DEP is in a "nondepressed" or "depressed" state. The implementation of DEP's automatic processing is based on insights afforded by a sophisticated, and biologically plausible, computer simulation of automatic and controlled processing (Schneider, 1987) and will be discussed when we describe DEP's analog of decreased motivation in depression. DEP can adapt to a changing environment by changing the weights connecting SS's nodes to each other and the environment, which is analogous to learning new emotional responses for directing attention.

PEAM is an associative network linking SS leaf nodes to SS inputs (Figure 2, PEAM). Using a simple learning rule, PEAM's weights are continuously adjusted until it can predict the environmental inputs that produce particular SS responses. Given an SS output, PEAM can reproduce the inputs to SS that should cause it. This allows PEAM to be used to retrain SS's responses using a model of the environment, instead of the environment itself. Retraining a self-schema with a model of the environment has two advantages. The model can select more relevant training inputs, and learning with a model is less dangerous than learning with the real thing. PEAM's associative structure can be used to trigger SS controlled processing and then to provide inputs most effective for modifying SS (discussed below).

Figure 3 displays six statistics calculated with respect to a DEP simulation run. A training epoch is one exposure to each possible environmental input. SS and PEAM's weights are initially set to small random numbers and adjusted with simple learning rules until SS produces correct outputs to environmental inputs E_0 and PEAM produces correct inputs to SS given the SS outputs (Figure 3, E_0).

DEP's analog of a *failure* (Figure 3, (1)) is the generation of an incorrect response to environmental inputs. We can precipitate a failure by abruptly changing DEP's environment from E_0 to E_1 (Figure 3, "Environmental Change"). This is analogous to a personal failure caused by an automatic misdirection of attention to inappropriate goals and plans. However, a single failure should not induce DEP to modify and reautomate its attention-directing weights. The failure must meet the same criteria that predict vulnerability to depression in humans. It must be *stable*, *internal*, and *global* (Figure 3, (2)(3)(4)). The failure must be likely to reoccur (stable failure). The failure must be avoidable by changing DEP's self-schema

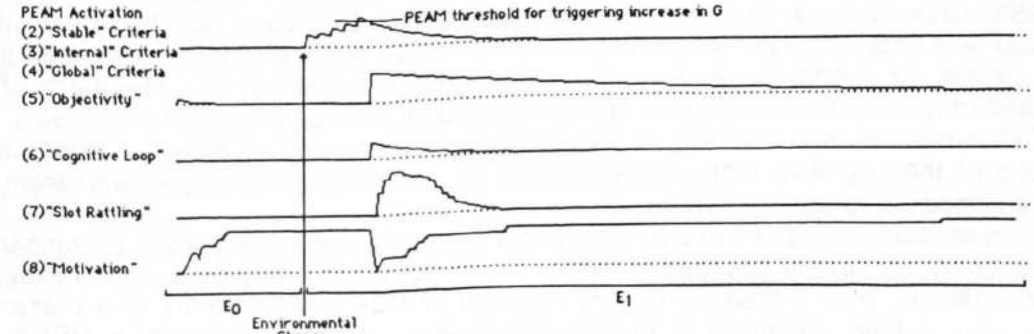

Figure 3: Output from DEP Simulation

(internal failure). The cause of the failure must potentially affect many different leaf nodes (global failure). All three of these criteria for triggering DEP's reactive depression can be achieved by using total PEAM activation to trigger controlled processing of DEP's self-schema.

PEAM accepts inputs and uses an interactive activation and competition paradigm (McClelland, 1981) to settle into an equilibrium state. PEAM's total activation depends on how much input it receives, where it receives the input, and how strongly individual nodes activate or suppress each other. Each time SS encounters a set of inputs and activates a leaf node, PEAM is activated with an input corresponding to SS's response. When SS is wrong and can find a correct response, the PEAM node associated with the correct response is also activated. If SS is wrong once (a non-stable failure), PEAM will not achieve a large activation because PEAM's weights have not had time to converge to a new model of the environment. If SS cannot find a correct response (a non-internal failure), PEAM will not achieve a large activation because PEAM does not receive the second input. If the two inputs (Figure 2, "False Negative Path" and "False Positive Path") are too similar (a non-global failure), PEAM will not achieve a large activation because not enough nodes are activated. Thus we can "gate" controlled processing of DEP's self-schema according to the criteria that a failure must be stable, internal, and global by establishing a threshold activation for DEP's analog of emotional memory. When PEAM activation (Figure 3, (2)(3)(4)) reaches threshold, DEP's depression commences. PEAM triggers DEP's depression by increasing a parameter (G) that increases the sensitivity of SS's weights to feedback from the environment about SS's performance.

DEP's analog for degree of *objectivity* (Figure 3, (5)) during reactive depression is the effect of a learning signal used to change SS weights based on the difference between SS's actual and desired output (a delta rule). Since increased sensitivity to differences between actual and desired performance is one kind of objectivity, increasing the effect of the learning signal has the effect of increasing DEP's analog of objectivity in the service of accurate self-schema modification. Activation of SS's nodes uses a standard logistic function (Rumelhart & McClelland, 1986), containing an additional parameter, G, divided into the net input and controlling the slope of the sigmoidal output of the logistic function. A small G results in extremely high or low SS node activation, and SS weights that are insensitive to the learning signal. A large G results in SS node activations near zero, and SS weights that are very sensitive to the learning signal. Decreasing G results in automatic descent of DEP's self-schema because all SS node activations are larger than a minimum threshold. Increasing G results in controlled descent because node activations fall below minimum thresholds and SS nodes must be sequentially searched to find the correct descent path. A small G is necessary for highly automated responses to a relatively static environment. A large G is necessary for learning new weight configurations in a changing environment. DEP's analog of emotional memory, PEAM, triggers modification and reautomation of DEP's self-schema by increasing G and allowing it to decay to zero. However, PEAM will only trigger controlled processing if the failure is stable, internal, and global, the same predictors for vulnerability to depression in humans.

DEP's analog of the *cognitive loop* (Figure 3, (6)) is the input from PEAM to SS providing training cases for modifying SS's weights (Figure 2, "Cognitive Loop"). Activating PEAM with inputs from SS's false positive and false negative leaf nodes creates a coalition of positively activated, similar leaf nodes. These are exactly the leaf nodes (along with their associated inputs to SS) most useful for modifying and reautomating DEP's self-schema. Members of this coalition are randomly selected to generate appropriate inputs and learning signals for SS . This sequence of failure-related inputs to SS, in the absence of actual inputs from the environment, is DEP's analog of the cognitive loop, the sequence of failure-related thoughts and memories preoccupying the mildly depressed person.

DEP's analog of *slot rattling* (Figure 3, (7)) is the increased weight variation accompanying a large learning signal and easily changed weights. Since slot rattling may be a sign of increased self-schema malleability, then increased weight variation in DEP's self-schema is a reasonable analog for the slot rattling observed in human depression. Weight variation in DEP's self-schema is especially severe if the environment is noisy and inconsistently rewards and punishes the same response.

DEP's analog of *lack of motivation* (Figure 3, (8)) is SS's lack of automaticity. Assuming "motivation" is how forcefully and persistently attention is directed to particular goals, motivation is how automatically goals are pursued. If an SS node (Figure 2, branch points in "Self-Schema") achieves activation above a minimum threshold, then it "automatically" transmits this activation, in parallel, to its daughter nodes at the next level down. If an activation falls below the minimum threshold, the node must wait for an external "controlling" signal to transmit. Increasing G increases feedback from the environment at the same time it causes all activations to move toward zero. As activations drop below minimum thresholds, more and more external controlling signals are necessary to achieve descent down DEP's self-schema. The external controlling signals (presumably requiring an "effortful" allocation of attention) are necessary to affect a sequential search for a correct descent path. As G decays back to zero, activations increase and descent of DEP's self-schema becomes more automatic, or "motivated". Thus, DEP must pass through a phase of controlled processing to learn new automatic processes. In order for DEP to change from one set of motivated, but incorrect, behaviors to a new set of equally motivated, and correct, behaviors, it must pass through a phase of unmotivated, but objective, modification and reautomation of its analog of the self-schema.

Conclusion

DEP is a highly simplified implementation of one way that emotions might normally direct attention to appropriate goals and plans, and periodically reorganize in the face of a changing environment. DEP exhibits a number of depressive information processing phenomena including: stable, internal, and global failures triggering qualitative changes in behavior, increased objectivity, the cognitive loop, fluctuations of self-describing adjectives, and decreased motivation. An emotional attention-directing system, grossly equated with the self-schema, normally operates in an automatic and motivated fashion insulated from environmental feedback. An emotional memory system constructs a model of the environment and periodically triggers reorganization of the emotional attention-directing system. Stable, internal, and global failures trigger the reorganization. These criteria follow from the self-schema's hierarchical structure, and are implemented using associative emotional memory. Once self-schema reorganization begins, emotional memory serves as a reservoir of training inputs to the self-schema. Loss of automaticity for engaging the environment causes a withdrawal that protects the vulnerable system while it reorganizes.

This computational model of reactive depression is one of many possible models. The DEP computer simulation forces us to examine our assumptions about what constitutes our theory. We need a more complete elaboration about what DEP's nodes and weights represent in terms of specific environmental stimuli, emotional processes, plans, and goals. One avenue is to represent a depressive scenario based on a case description of reactive depression. Another possibility is to relate DEP's parameters and behavior to some of the many standardized tests for diagnosing depression. Eventually it may be possible to relate DEP subsystems and processes to neurologic changes during depression. A rapidly evolving literature on the "neurology of depression" (Otto et al., 1987; Coffey, 1987) implicates the right cerebral hemisphere in depression. The right hemisphere plays a special, if not dominant, role in emotional, attentional, and pragmatic processes (Code, 1987; Kosslyn, 1987). Since our theory of reactive depression is about these topics, our computational model may provide a unique perspective to analyze previously disparate observations about the role of the right hemisphere in depression. Eventually we may be able to advance specific hypotheses about differences between mild-reactive and chronic-suicidal depression by understanding the differences in terms of perturbations in a cognitive model of reactive depression (Webster et al., 1988).

References

Alloy, L., & Abramson, L. (1979). Judgement of contingency in depressed and nondepressed students: Sadder but wiser. *Journal of experimental psychology*: General, 108, 441-485.

Beck, A., Rush, A., Shaw, B., & Emery G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Blaney, P. (1986). Affect and memory: A review. *Psychological bulletin*, 99, 229-246.

Bower, G. (1981). Mood and memory. *American psychologist*, 36, 129-148.

Carver, C. & Scheier, M. (1986). Functional and dysfunctional responses to anxiety: The interaction between expectancies and self-focused attention. In R. Schwarzer (Ed.) *Self-related cognitions in anxiety and motivation.* Hillsdale, Lawrence Erlbaum.

Code, C. (1987). *Language, aphasia, and the right hemisphere.* New York: John Wiley & Sons.

Coffey, C. (1987). Cerebral laterality and emotion: The neurology of depression. *Comprehensive psychiatry*, 28, 197-219.

Coyne, J., & Gotlib, I. (1983). The role of cognition in depression: A critical appraisal. *Psychological Bulletin*, Vol. 94, 3, 472-505.

de Sousa, R. (1987). *The rationality of emotion.* Cambridge, MIT Press.

Dyer, M. (1983). *In-depth understanding.* Cambridge: MIT Press.

Feigenbaum, E. & Simon, H. (1984). EPAM-like models of recognition and learning. *Cognitive Science*, 8, 305-336.

Giles, D., & Shaw, B. (1987). Beck's cognitive theory of depression: Convergence of constructs. *Comprehensive psychiatry*, 28, 416-427.

Ingram, R. (1984). Toward an information-processing analysis of depression. *Cognitive therapy and research*, 8, 443-478.

Klerman, G. (1987). The nature of depression: mood, symptom, disorder. In A. Marsella, R. Hirschfeld, & M. Katz (Eds.), *The measurement of depression.* New York: Guilford Press.

Kosslyn, S. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological review*, 94, 148-175.

Marcus, H. (1980). The self in thought and memory. In D. Wegner and R. Vallacher (Eds.), *The self in social psychology.* New York: Oxford University Press.

McClelland, J. (1981). Retrieving general and specific knowledge from stored knowledge of specifics. *Proc. of the third annual meeting of the cognitive science society*, 170-172.

Otto, M., Yeo, R., & Dougher, M. (1987). Right hemisphere involvement in depression: Toward a neuropsychological theory of negative affective experiences. *Biological psychiatry*, 22, 1201-1215.

Ross, M. (1985). Depression, self-concept, and personal constructs. In F. Epting & A. Landfield (Eds.), *Anticipating personal construct psychology.* Lincoln: University of Nebraska Press.

Rumelhart, D., McClelland, J. (1986) *Parallel distributed processing, Explorations in the microstructure of cognition, Vol. 1: Foundations* . Cambridge: MIT Press.

Schneider, W., Dumais, S., & Shiffrin, R. (1984). Automatic and control processing and attention. In *Varieties of attention.* New York: Academic Press.

Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, 84, 1-66.

Schneider, W. & Detweiler, M. (1987). A connectionist/control architecture for working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21), New York: Academy Press.

Seligman, M. (1984). Causal explanations as a risk factor for depression: Theory and evidence. *Psychological review*, 91, 347-374.

Shavelson, R. & Marsh, H. (1986). On the structure of the self-concept. In R. Schwarzer (Ed.) *Self-related cognitions in anxiety and motivation.* Hillsdale, Lawrence Erlbaum.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological review*, 84, 127-190.

Simon, H. (1967). Motivational and emotional controls of cognition. *Psychological review*, 74, 29-39.

Webster, C., Sherman, H., & Banks, G. (1988). Computational neurology: The perturbation of normal models of cognition. *Proceedings of the AAAI symposium on artificial intelligence in medicine*, Stanford Univeristy, Palo Alto, CA.