# Technical Report on “Facial Expression Analysis in Schizophrenia”

Excerpts taken from:

Carter, J.R. (2000). Facial expression analysis in schizophrenia. Unpublished Ph.D. dissertation, Department of Psychology, University of Western Ontario, London, Ontario, Canada.

Pages 56 - 58

## Conversion to Real Time

One method that connectionist researchers have used to represent latency is number of iterations (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996). Ratcliff, van Zandt, and McKoon (1999) included “some parameters to scale number of iterations to time” (p. 279), but did not elaborate on how the parameter values were determined. An important contribution of the current project is a formula derived by closed-form solution to enable the model to make predictions in real time. This goal was accomplished by borrowing from Busemeyer and Townsend’s (1993) model the idea of a parameter, *h*, representing the amount of real time that passes during each time step (iteration). Busemeyer and Townsend did not provide a specific formula.

To develop a formula, the latency component of the *SS* formula was selected as a lack-of-fit index because it is commonly used, conceptually simple, and easy to calculate:

$$\sum_{i=1}^{N} (h\,\mathrm{K}_i - RT_i)^2.$$

The following formulae (all calculated in MAPLE V.4, Symbolic Computation Group, 1995) determine the value of *h*. The minimum $SS_{latency}$ with respect to *h* can be represented as

$$MIN_h\Big[\sum_{i=1}^{N} (h\,\mathrm{K}_i - RT_i)^2\Big] = \sum_{i=1}^{N} \frac{\Big[\mathrm{K}_i\big(\sum_{j=1}^{N} \mathrm{K}_j RT_j\big) - RT_i\big(\sum_{j=1}^{N} \mathrm{K}_j^2\big)\Big]^2}{\big(\sum_{j=1}^{N} \mathrm{K}_j^2\big)^2}.$$

That is, the left-hand side of the equation, the minimum value for the sum-of-squares error function – assuming an optimum value for *h* – can be calculated according to the function on the right-hand side of the equation. Subsequently ignoring the minimization operator and isolating *h* yields the elegantly simple equation,

$$h = \frac{\sum_{i=1}^{N} (\mathrm{K}_i \, RT_i)}{\sum_{i=1}^{N} \mathrm{K}_i^2}.$$

Alternately, the function for *h* can be derived by differentiating the *SS* function with respect to *h*, setting that function to zero, and solving for *h*. The minimum value of the *SS* function occurs when its slope equals zero. That is,

$$0 = \sum_{i=1}^{N} (2\mathrm{K}_i(h\mathrm{K}_i - RT_i)).$$

Solving this equation for *h* yields the same function as before. That *SS* is minimized (rather than maximized) is confirmed because the second derivative of *SS* with respect to *h* is always positive,

$$\sum_{i=1}^{N} (2\,\mathrm{K}_i^2).$$

# CHAPTER FOUR: SCHEMATIC FACES STUDY

The first study examines extant data concerning the content and latency of similarity judgments of schematic faces. In previous work (Carter, 1994; Carter & Neufeld, 1996a, 1996b, 1999), the content of judgments were subjected to a multidimensional-scaling analysis. The resulting parameters (specifically, the mean weight assigned to each dimension for each group) were successfully predicted by mean response latencies for subsequent dimensions. The current analysis takes a different tack by modeling each response and corresponding latency. A connectionist simulation was created and damaged. This analysis is qualitatively different from earlier work in that the previous analysis was primarily computational and the current analysis is more algorithmic (see Marr, 1982, p. 25). This small simulation not only paves the way for the second study, which uses pictures of real faces (with many more potential features), but it also demonstrates the effectiveness of the formulae for calculating the real time scaling parameter, *h*, and the lack-of-fit criterion "$\chi^2$."

## Method

### Participants, Procedure, and Stimuli

Characteristics of the participants and research method are described in detail in previous work (e.g., Carter, 1994; Carter & Neufeld, 1999). Participants were 20 each of nonpatients,

patients with paranoid schizophrenia, and patients with other forms of schizophrenia. Forty-seven were men, and 13 were women. The mean ages (and standard deviations) for the three groups, respectively, were 30.55 (10.13), 37.95 (9.26), and 42.3 (10.74) years. Groups were approximately matched in terms of handedness, vocabulary, and social class. Each participant was tested individually in two sessions. The first session consisted of paper and pencil questionnaires and a diagnostic interview. The second session was devoted to making complex (i.e., global) and simple (i.e., on a single dimension) similarity judgments regarding pairs of words and pairs of schematic faces. Only the global judgments of schematic faces pertain to the current study. Participants were presented with all unique pairs of stimuli within each class. Figure 5 depicts the schematic faces (which were taken from Lay, Waters, & Park, 1989) as they vary on two dimensions: Arousal and Pleasure. Participants rated the difference in the meaning of the expressions for each pair of faces on a nine-point scale ranging from *Highly Similar* (1) to *Highly Dissimilar* (9). Their responses were timed. The response latencies were adjusted to compensate for physical response time (assumed to be 0.16 s), and to eliminate responses presumed to represent fast guessing (less than 0.1 s), and derailment (greater than 30 s). No trials were eliminated for fast guessing, but about one per cent of all judgments were eliminated for derailment.

## Simulation

### Architecture

Several architectures for the backpropagation network were tested. A 24-unit model, depicted in Figure 6, was selected on the basis of adequacy of fit to data and simplicity. The input units connected to a set of five hidden units. The hidden units were connected to a single output unit, which also sent activation to the hidden units. Thesame architecture was used in two separate training runs.

### Representations

The input units were divided into two sets of nine. Each “face” was equally likely to be presented on the “left” (i.e., the first nine input units) or “right” (i.e., the second nine input units). Within each set, two active units (i.e., activation values set to one) represented each face. One active unit was drawn from a subset of six to represent the shape of the mouth. The other active unit was drawn from a subset of three to represent the shape of the eyes. The remaining input units’ activations were set to zero. Model predictions of dissimilarity were represented as the activation of the output unit. Low activations signified that the two stimuli were similar, and high activations signified that they were different.

### Learning Rule and Associated Parameters

The network was trained with the backpropagation-through-time algorithm (as described in Appendix A, and instantiated on Carter, 1997 - 1999). Weights were updated at the end of each training epoch. Each epoch consisted of the entire set of unique pairs of faces, totaling 36 patterns, each presented once. The learning rate, $\varepsilon$, was initially set to one, and momentum values for both weights and changes in weights were initially set to zero. At 44, 864 epochs, the second run became unstable because the learning rate was too great. Therefore, training was continued with ε reduced to .1 and momentum for weights increased to .5.

## Processing of a Pattern

Input activations were held constant throughout each trial. Randomness can actually increase the effectiveness of nonlinear, dynamic systems (Rossi, 1996). During training, therefore, the hidden and output unit activations were set at the beginning of each trial according to a random (rectangular) distribution with a mean of .2 and a standard deviation of .025. During testing, however, the starting activations were always exactly .2, to ensure that the lack-of-fit indices were reproducible whenever possible. Processing continued for each trial until the network settled.[1] Specifically, the change in activation (excluding any added noise) for all units from one iteration to the next had to be less than an arbitrarily small value, ι. Within each iteration, all layers were updated sequentially, beginning with input, followed by hidden, and ending with output.

## Training Regime

The initial weights were drawn from a random (rectangular) distribution with a mean of zero and a standard deviation of 0.15. Because attempts to estimate the parameters associated with time before an approximate fit for content was achieved failed, the network was trained to a total backpropagation error (see Appendix A) of less than 0.3 (an arbitrary value that worked) before those parameters were introduced. That is, initially, τ was set to one and each trial continued for exactly three iterations. To facilitate an adequate fit, τ and ι were set at their optimum values (as determined by the line-search parameter estimation algorithm) as soon as the network had achieved some degree of accuracy (i.e., backpropagation error less than 0.3, which occurred at 3, 785 epochs for the first run and at 3, 778 epochs for the second). The training data (see below) was used for all of these calculations, preventing the loss of any degrees of freedom. Training continued until the network achieved an adequate fit to the observed data, as defined by a Maximum Likelihood criterion. Given the structure of the mathematical model and the data at hand, the parameter values that lead to the modal value of the error distribution are the most likely values. The modal value of the $\chi^2$ distribution (and, therefore, theoretically, the "$\chi^2$" distribution) occurs at *df* - 2 (in this case, 69).

## Basic Analyses

The data were collapsed into four sets of responses. That is, four "homogeneous subjects" (Townsend, 1984, pp. 386 - 387) were constructed, two from the nonpatient group, and one from each of the schizophrenia groups. Having two homogeneous subjects constructed from nonpatient data allowed for cross-validation of the model (see Browne, 2000). One set of

---

[1] Processing was also limited to an arbitrary maximum of 250 iterations within a trial.

responses was used for determining the parameter values (including weights and biases), and then the adequacy of the fit was tested against the other. Each set consisted of 36 content and latency pairs, one pair for each possible judgment. Content consisted of the mean ratings on the nine-point scale (scaled to fall between zero and one by dividing by ten), and latency was measured in seconds and adjusted as previously described. A second run of the simulation reversed the nonpatient training and testing data sets and had a different set of random starting weights.

After an adequate fit to the control data was achieved, the model was damaged to simulate patient data. All of the types of damage described in Chapter Two were pitted against each other in direct competition. In all cases, parameter values were estimated using the test data, resulting in the loss of one degree of freedom for each parameter estimated. This procedure allowed for a quantitative test of any improvement in fit. Conceptually, the undamaged model is the special case of (i.e., "is nested in") the damaged model where the damage parameter is set to a value such that it has no effect (e.g., multiplying by one or cutting zero units). Mathematically, "$\chi^2_{df}$" - "$\chi^2_{df-m}$" = "$\chi^2_{m}$", where *df* is the number of degrees of freedom in the undamaged model, and *m* is the number of parameters estimated. Similarly, a model is nested in a model with one or more parameter values determined by some other set of data. (See Bamber & van Santen, 1985, for a discussion of nested-nesting relations.)

## General Principles for Evaluating Models

Many considerations – not one of which is absolute – contribute to the evaluation of various competing explanations (McCracken, 1988). Those interventions that not only improve fit but also reproduce empirical observations are obviously superior to those that do not. Improvements that are not "statistically significant"[2] (i.e., "$p$" > .1, *NS*) are not reported. Although large improvements are more impressive than small ones, care was taken not to accept a model with an artificially low lack-of-fit index (as might be indicated by a significantly-negative "$z$"). Further, the model with the closest approximation to the data is not necessarily the best (Zucchini, 2000). In the current project, consistency in parameter values was weighted heavily (Meehl, 1983; Meehl & Golden, 1982). Interventions that were replicated across cross-validation runs, with different fit criteria, and across tasks are stronger contenders than those that were less consistently beneficial. Another facet of consistency is the identifiability (in the mathematical sense) of the parameter values. A less obvious principle is that positive and negative findings are not equal. Those cases when a model is rejected should carry more weight than when it is tentatively accepted (Meehl, 1978). The cost of incorrectly rejecting a valid model is far greater than the cost of tentatively accepting an invalid model for future consideration. Therefore, each test was treated as an independent hypothesis with no correction for Type I error. Other factors are parsimony and specificity. Another principle is that a good model directs, or even inspires, future research. A final consideration is the extent to which current findings are compatible with extant literature.

---

[2] Recall that the assumptions regarding the "$\chi^2$" distribution are quite restrictive, potentially limiting the veracity of exact significance tests.

## Results and Discussion

### Unperturbed Networks

The simulation adequately reproduced the data collected from the control group. The values of $\tau$ and $\iota$ were fixed at their estimated values of .095 and .014, respectively.[3] By epoch 91, 177, "$\chi^2_{71}$" was 70.54 and $h$ was 0.322873. The simulation is considered valid and an adequate fit to the data.[4] Table 2 presents the human and model values for both similarity and latency. Figure 7 compares the two cross-validation samples and the model in terms of similarity and latency. Ideally, the best-fitting regression line would pass through the origin and have a slope of one, and the data would lie close to the line. Note that the model and the cross-validation sample are comparable in terms of similarity judgments. The model performed less well with respect to latency than similarity. It appeared, however, to outperform the cross-validation sample, which can occur with noisy data because the model contains no element corresponding to measurement error.

### Data Collected from People with Paranoid Schizophrenia

The simulations were tested against the data from patients with paranoid schizophrenia using the control group's value for $h$ and no damage. The fit was unacceptable. On the first run, "$\chi^2_{72}$" was 97.58 ("$p$" $\approx$ .025). Recalculating $h$ did not improve the fit ("$\chi^2_{71}$" = 96.27, "$p$" $\approx$ .025; $\Delta$"$\chi^2_1$" = 1.31, *NS*). Therefore, the value for $h$ from the control data was used for further analyses. For the second run,"$\chi^2_{72}$" was 155.81 ("$p$" < .001), but applying the value for $h$ based the paranoid data, 0.314843, improved "$\chi^2_{71}$" to 87.66 ("$p$" $\approx$ .1; $\Delta$"$\chi^2_1$" = 68.15, "$p$" < .001). Therefore, for the paranoid data, the first run is compared to "$\chi^2_{72}$" of 97.58, with the control data value for $h$, and the second run is compared to "$\chi^2_{71}$" of 87.66, with a recalculated value for $h$ (unless otherwise noted).

---

[3] The results of the first run only are reported when the two runs yielded essentially the same pattern of results. Carter, 2000, reports on all significant results.

[4] "$z$" = -0.003, *NS*; $\chi^2_{69} = 5.15$, *NS*, $z = -1.52$, $.1 > p > .05$; *model-mean*-$\chi^2_{71} = 93.3$, $p \approx .5$, $z > 0$, *NS*.

Table 2

Human and model predictions of dissimilarity ratings (1..9) and latencies (s) for the first run of the schematic faces network.

| | Human | | Model | |
|---|---|---|---|---|
| Stimulus Pair | Dissimilarity | Latency | Dissimilarity | Latency |
| 1, 2 | 4.18 | 5.23 | 4.46 | 6.46 |
| 1, 3 | 6.55 | 5.18 | 6.53 | 7.1 |
| 1, 4 | 3.46 | 5.42 | 4.84 | 4.84 |
| 1, 5 | 6.64 | 4.94 | 6.58 | 6.78 |
| 1, 6 | 8.27 | 7.81 | 6.46 | 4.84 |
| 1, 7 | 3.64 | 5.5 | 4.46 | 8.07 |
| 1, 8 | 6.36 | 6.1 | 6.68 | 7.1 |
| 1, 9 | 8.36 | 7.87 | 7.26 | 6.78 |
| 2, 3 | 3.82 | 5.34 | 5.8 | 5.17 |
| 2, 4 | 6.27 | 7.61 | 6.93 | 7.1 |
| 2, 5 | 3 | 6.36 | 4.17 | 4.2 |
| 2, 6 | 5.82 | 7.12 | 7.54 | 6.78 |
| 2, 7 | 6.09 | 6.5 | 5.98 | 3.23 |
| 2, 8 | 7.91 | 8.11 | 7.3 | 8.39 |
| 2, 9 | 6.36 | 5.52 | 6.79 | 6.13 |
| 3, 4 | 7.64 | 5.88 | 6.81 | 7.43 |
| 3, 5 | 6.36 | 7.64 | 6.2 | 4.52 |
| 3, 6 | 3.73 | 6.07 | 5.17 | 5.49 |
| 3, 7 | 8.46 | 6.69 | 6.76 | 8.07 |
| 3, 8 | 6.82 | 6.09 | 6.62 | 3.55 |

| | | | | |
|---|---|---|---|---|
| 3, 9 | 4.36 | 6.7 | 6.54 | 4.84 |
| 4, 5 | 4.73 | 5.53 | 4.46 | 2.58 |
| 4, 6 | 7.09 | 6.85 | 6.06 | 6.46 |
| 4, 7 | 2.91 | 6.93 | 2.73 | 6.78 |
| 4, 8 | 6.18 | 7.77 | 6.71 | 5.17 |
| 4, 9 | 7.18 | 7.97 | 6.74 | 3.55 |
| 5, 6 | 4.46 | 6.32 | 5.03 | 8.07 |
| 5, 7 | 7.18 | 7.49 | 6.54 | 8.72 |
| 5, 8 | 6.18 | 9.9 | 7.21 | 9.04 |
| 5, 9 | 6.18 | 8.88 | 6.77 | 6.46 |
| 6, 7 | 7.82 | 5.92 | 7.41 | 9.04 |
| 6, 8 | 6.36 | 8.69 | 7.79 | 9.36 |
| 6, 9 | 3.46 | 8.88 | 4.78 | 6.46 |
| 7, 8 | 4.27 | 6.2 | 5.03 | 5.81 |
| 7, 9 | 6.55 | 7.51 | 6.58 | 8.07 |
| 8, 9 | 3.91 | 6.05 | 4.45 | 7.75 |

### Data Collected from People with Nonparanoid Schizophrenia

The simulations were tested against the data from the patients with nonparanoid schizophrenia using the control group’s value for $h$ and no damage. For the first run, “$\chi^2_{72}$” was 83.8 (*NS*; $\chi^2_{70} = 50.8$, *NS*; $z = -1.71$, $p < .05$).[5] Calculating $h$ again did not improve “$\chi^2$.” Therefore, the value for $h$ from the control data was used for further analyses of the nonparanoid data on the first run. For the second run, “$\chi^2_{72}$” was 137.21 (“$p$” < .001; $\chi^2_{70} = 51.14$, *NS*; $z = -1.68$, $p < .05$), but applying the value for $h$ calculated from this data, 0.313744, improved this fit to a “$\chi^2_{71}$” of 87 (.1 > “$p$” > .05 ; Δ“$\chi^2_1$” = 50.21, “$p$” < .001). Applying the value for $h$ calculated for the paranoid data on the second run, 0.314843, yielded a similar result (“$\chi^2_{72}$” = 86.59, .1 > “$p$” > .05). For the nonparanoid data, the first run is compared to “$\chi^2_{72}$” of 83.8, with the control data value for $h$, and the second run is compared to “$\chi^2_{71}$” of 87, with a recalculated value for $h$ (unless otherwise noted).

## Additional Processing Hypothesis

Simply adjusting the settling criterion led to an improvement in fit for both patient data sets. Specifically, for the paranoid data, setting $\iota$ to .0093 led to “$\chi^2_{71}$” of 92.97 (“$p$” ≈ .05; Δ“$\chi^2_1$” = 4.61, “$p$” < .05). The resulting model is both necessary and sufficient by orthodox significance tests.[6] As predicted by a model in which the people with paranoid schizophrenia introject additional steps, the value of $\iota$ is smaller for the patient group than for the control group.

Adjusting the settling criterion also improved the fit of the model to the nonparanoid data. The optimum value for $\iota$ on the first run was .0136 (“$\chi^2_{71}$” = 72.23, *NS*; Δ“$\chi^2_1$” = 11.57, “$p$” < .05). Using the value estimated from the paranoid group’s data was significantly detrimental to the fit (“$\chi^2_{72}$” = 88.56, .1 < “$p$” < .05; Δ“$\chi^2_1$” = 16.33, “$p$” < .001). On the second run, the best fit was found with $\iota$ equal to .0092, regardless of whether $h$ was calculated from the nonparanoid group data or the control group data.[7] The model with $\iota$ and the control group value for $h$ is both necessary and sufficient by orthodox significance test.[8] In both cases, substituting the value for $\iota$ estimated from the paranoid group was as effective as estimating a value from the nonparanoid group in terms of “$\chi^2$.” In contrast, the paranoid value for $\iota$ (with the control group $h$) was clearly inappropriate for the nonparanoid group data from the perspective of orthodox significance testing (*model-mean*-$\chi^2_{72} = 1,019.1$, $p < .001$; $\Delta\chi^2_1 = 978.39$, $p < .001$).

---

[5] *model-mean*-$\chi^2_{72} = 50.99$, *NS*, $z = -1.86$, $p < .05$.

[6] $\chi^2_{69} = 54.34$, *NS*; $z = -1.19$, *NS*; $\Delta\chi^2_1 = 15.99$, $p < .001$; *model-mean*-$\chi^2_{71} = 68.78$, *NS*; $z = -0.06$, *NS*; $\Delta\chi^2_1 = 4.68$, $p < .05$.

[7] “$\chi^2_{70}$” = 83.78, *NS*; Δ“$\chi^2_1$” = 3.22, .1 > “$p$” > .05; “$\chi^2_{71}$” = 79.65, *NS*; Δ“$\chi^2_1$” = 57.56, “$p$” < .001.

[8] $\chi^2_{69} = 39.46$, *NS*; $z = -2.82$, $p < .01$; $\Delta\chi^2_1 = 11.68$, $p < .001$; *model-mean*-$\chi^2_{71} = 40.71$, *NS*; $z = -2.85$, $p < .01$; $\Delta\chi^2_1 = 51.77$, $p < .001$.

This finding is consistent with previous research suggesting that deficits among people with schizophrenia are caused by additional processing steps (Carter & Neufeld, 1999; Neufeld et al., 1993). The results of this study are not clear regarding whether different parameter values are needed to simulate paranoid and nonparanoid schizophrenia. At this point, the biological correlate for this parameter is unknown.

### Neurotransmitter Dysfunction Hypothesis

Changes in *gain* resulted in reduced “$\chi^2$” for both patient groups. This parameter is associated with the dysfunctional neurotransmitter hypothesis because both neurotranmitters and *gain* appear to influence the signal-to-noise ratio. In spite of the improvement in fit, the neurotransmitter dysfunction hypothesis was rejected for this task because the *gain* parameter lacked specificity and identifiability. As well, the fit may have been somewhat artifactual.

For the paranoid group data, the estimated value was 1.57 (“$\chi^2_{71}$” = 63.79, *NS*; “*z*” = -0.58, *NS*; Δ“$\chi^2_1$” = 33.79, “*p*” < .001). The *gain* parameter lacked identifiability, however. The best-fitting parameter values based on the alternate parameter estimation criteria (*SS* and $\hat{\chi}^2$) tended to be close to one. Further, an increased value for *gain* contrasts with Cohen and Servan-Schreiber’s (1992) results. One possibility is that the tasks were sufficiently different that they rely on different neurological substrates. As noted in Chapter Two, biologically-based arguments can be made for either increased or decreased dopamine in schizophrenia. Conceivably, different brain areas may have different maladaptive levels of neurotransmitters and also form the neurological substrate for different tasks.

Poor specificity also tends to argue against the *gain* parameter. When *gain* was adjusted for the control group, fit (as measured by “$\chi^2$”) improved significantly. This finding may suggest that the improvements resulting from changes of the *gain* parameter on this task are a property of the networks rather than a reflection of a legitimate biological correlate of schizophrenic deficits. A model artifact related to *gain* is a reasonable possibility. The model response is an interpretation of the activation of the single output unit. The activation function is sharpened by increasing the *gain* parameter. A sharper activation function may allow for less precise, but more variable, output values in that larger changes in afferent activations would be required for changes in output unit activation. These changes may affect both actual model accuracy and the test statistic.

Another possibility is that the fit is an artifact of the “$\chi^2$” criterion. Indeed, the alternate parameter estimation criteria (*SS* and $\hat{\chi}^2$) both became larger when *gain* was increased from one to the value yielding the lowest “$\chi^2$.” Further, the model-predicted variances were inflated. The observed standard deviations were 1.0828 for latency and 0.1934 for accuracy, but the model values were 2.4394 and 0.2436, respectively. This increased variability was reflected by certain model predictions for similarity judgments that exceeded the range of the human data. The reduction in “$\chi^2$” was not entirely artificial, however. As shown in Figure 8, individual trial predictions based on increased *gain* were somewhat better than predictions based on decreased $\iota$ with respect to both similarity ratings and latency.

Similarly, increasing *gain* also resulted in reduced “$\chi^2$” for the nonparanoid data. Similar additional analyses were completed. The same conclusions resulted – improvement related to *gain* was not specific to the patient group, and values were not identifiable. The decrease in lack-of-fit may to be an artifact of the criterion. Specifically, the model tends to overestimate the variances. Therefore, the neurotransmitter hypothesis was not supported for either patient group in this study.

## Cortical Pruning Hypothesis

Three type of interventions were applied to simulate cortical pruning: Reporting weights as zero, adding random noise to the activations, and multiplying by random numbers. The results were somewhat different depending on the patient group.

### Data Collected from People with Paranoid Schizophrenia

Neural Darwinist hypotheses. Reporting weights as zero did not improve fit on the first run. On the second run, however, eliminating 39 per cent of the weights – specifically, those with the lowest squared values, $\partial E / \partial w$, and $(\partial E / \partial w)^2$ – and substituting the control value for *h* led to a “$\chi^2_{71}$” of 142 (“$p$” < .001; Δ“$\chi^2_1$” = 13.81, “$p$” < .001). This improvement was also found after applying a recalculated value for *h* (“$\chi^2_{70}$” = 83.77, *NS*; Δ“$\chi^2_1$” = 3.89, “$p$” < .05, as compared to the undamaged network; $\Delta \chi^2_1 = 58.23$, “$p$” < .001, as compared to the damaged network with the control group’s *h* value). This result provides limited support for the hypothesis that the least important weights are systematically eliminated, possibly because of an exaggerated developmental pruning process.

Inhibitory versus excitatory synaptic elimination. Adding noise to the hidden units tended to improve fit. Because the estimated mean was zero, the standard deviation was estimated again by itself and found to be 0.167 for a rectangular distribution (“$\chi^2_{71}$” = 73.68 [S.D. = 5.9501], *NS*; Δ“$\chi^2_1$” = 23.9, “$p$” < .001). The results for second run and the Gaussian distribution on both runs were all similar. They strongly support a model in which synaptic connectivity is disrupted, but the type of synapse (excitatory or inhibitory) is irrelevant.

Underconnectivity versus overconnectivity. An anomaly related to synaptic density was also supported because multiplying each weight by a random number also tended to improve fit. The nature of the anomaly was unclear, however. When the distortion occurred every tick and the distribution was rectangular, the mean was one and the standard deviation was 0.393 (“$\chi^2_{71}$” = 83.87 [S.D. = 6.5647], *NS*; Δ“$\chi^2_1$” = 13.71, “$p$” < .001). In contrast, when the distribution was Gaussian, the best fitting mean and standard deviation, respectively, were 2.5 and 0.667 (“$\chi^2_{70}$” = 67.45 [S.D. = 4.0738], *NS*; “$z$” = -0.18, *NS*; Δ“$\chi^2_2$” = 30.13, “$p$” < .001). The results for the second run and when the distortion occurred every trial for both runs were all similar. Because the average strength of connections remained the same for the rectangular distribution, the hypothesis that the wrong synapses are eliminated was somewhat supported. In contrast, the increased mean with the Gaussian distribution could suggest overconnectivity.

## Data Collected from People with Nonparanoid Schizophrenia

The cortical pruning hypotheses did not fair as well with the data from the nonparanoid group. Cutting weights, adding activation noise, and multiplying the weights all yielded inconsistent results.

Neural Darwinist hypotheses. On the first run, eliminating the one per cent of weights with the highest values seemed effective (“$\chi^2_{71}$” = 67.57, *NS*; “*z*” = -0.25, *NS*; Δ“$\chi^2_1$” = 16.23, “*p*” < .001). In contrast, on the second run, eliminating the 63 per cent of weights with the lowest squared, $\partial E / \partial w$, and $(\partial E / \partial w)^2$ values led to a trend towards improvement (“$\chi^2_{70}$” = 84.25, *NS*; Δ“$\chi^2_1$” = 2.75, .1 > “*p*” > .05). These results provide weak evidence for the two conflicting hypotheses that the most and least important synapses are eliminated.

Inhibitory versus excitatory synaptic elimination. On the first run, adding noise to the hidden units from a rectangular distribution with a standard deviation of 0.138 each tick improved fit (“$\chi^2_{71}$” = 79.5449 [S.D. = 6.5714], *NS*; Δ“$\chi^2_1$” = 4.26, “*p*” < .05). The result was similar when noise was recalculated only once per trial. These findings were not replicated on the second run, providing only weak support for the hypothesis that damage occurs and is independent of the type of synapse (inhibitory or excitatory).

Underconnectivity and overconnectivity. Multiplying the weights by random numbers improved fit on the first run, but only trends towards improvement were found for the second. Resetting these values each tick drawing from a rectangular distribution with a mean of 2.251 and a standard deviation of 0.502 improved fit on the first run (“$\chi^2_{70}$” = 66.93 [S.D. = 3.3324], *NS*; “*z*” = -0.22, *NS*; Δ“$\chi^2_2$” = 16.87, “*p*” < .001). Applying the parameter values estimated from the paranoid data was detrimental to the fit (“$\chi^2_{72}$” = 81.27 [S.D. = 4.4799], *NS*; Δ“$\chi^2_2$” = 14.34, “*p*” < .001). Only a trend towards improvement was found on the second run. The results for the Gaussian distribution were similar.

Applying the distortion each trial sometimes improved the fit. On the first run, a rectangular distribution with a mean of 1.687 and standard deviation of 0.331 led to a “$\chi^2_{70}$” of 68.58 (S.D. = 8.58; *NS*; “*z*” = -0.08, *NS*; Δ“$\chi^2_2$” = 16.22, “*p*” < .001). This finding was not replicated on the second run. Applying the values estimated on the paranoid data did not significantly change the fit (“$\chi^2_{72}$” = 72.57 [S.D. = 6.5955], *NS*; Δ“$\chi^2_2$” = 3.98, *NS*). These results were replicated for the Gaussian distribution. These simulations provide weak evidence for the hypothesis that too few synapses are eliminated.

## Comment

The cortical pruning hypothesis seems better suited to account for the data on this task for the paranoid schizophrenia group than for the nonparanoid schizophrenia group. In particular, results were consistently replicated for the paranoid data but not for the nonparanoid data. This finding is somewhat surprising because structural damage is more often associated with nonparanoid symptoms. At the same time, however, the results may speak to the relative subtly of cortical pruning as compared to neural atrophy. The particular pattern of deficits is also important. Random elimination of weights was never effective in improving fit, and improvement by eliminating the most important weights was not replicated. Overall, the hypothesis that too many synapses are eliminated (or fail to develop) in schizophrenia appears more likely than the alternate hypotheses that too few or the wrong synapses are eliminated (or fail to develop).

## Neural Atrophy Hypothesis

For the data collected from people with paranoid schizophrenia, fit improved significantly when a proportion of units with the highest sums of afferent weights were rendered inactive (i.e., activation always reported as zero). In all cases, the output unit was protected from elimination. The estimated value was .14 (“$\chi^2_{70}$” = 92.76, *NS*; Δ“$\chi^2_1$” = 4.82, “*p*” < .05). Eliminating units by other criteria failed to produce a significant improvement in fit.

For the data collected from people with nonparanoid schizohrenia, eliminating the 12 per cent of the units with the greatest sum of afferent weights gave a “$\chi^2_{71}$” of 80.2 (*NS*; Δ“$\chi^2_1$” = 3.6, .1 > “*p*” > .05). This finding was only replicated on the second run when the nonpatient, training value was applied for *h*. In that case, the proportion of units eliminated was estimated at 9 per cent (“$\chi^2_{71}$” = 129.82, “*p*” < .001; Δ“$\chi^2_1$” = 7.39, “*p*” < .01). Applying the values estimated from the paranoid data was detrimental to the fit for the first run (“$\chi^2_{72}$” = 179.56, “*p*” < .001; Δ“$\chi^2_1$” = 99.36, “*p*” < .001), but made no significant difference for the second.

These results tend to support the idea that the well-documented neural atrophy found in schizophrenia may play a role in this similarity judgment task. Of note, however, random elimination was not effective. Of the several criteria for operationalizing importance, only the greatest sum of afferent weights resulted in schizophrenic responding. This result suggests that perhaps damage to relatively few – but critical – neurons impairs functioning on this task. The exact location of these neurons cannot be determined from this simulation, but two likely possibilities are the cingulate gyrus and frontal lobes. These areas are known to be implicated in both schizophrenia and facial affect perception (Deakin, 1994). It is unclear from the results of this study whether the same proportion of damage is responsible for deficits among people with paranoid and nonparanoid schizophrenia.

## Early Processing Hypothesis

Adding noise to the input units was usually detrimental to the fit. Estimated values for the mean and standard deviation were typically both zero. One exception to this finding was that adding noise drawn from a rectangular distribution with a mean of 0.127 and a standard deviation of 0.219 every tick improved the fit for the second run (“$\chi^2_{69}$” = 79.33 [S.D. = 5.9474], *NS*; Δ“$\chi^2_2$” = 8.33, “*p*” < .025). Similar results were found with a Gaussian distribution. This finding provides only weak evidence for the possibility that early visual processing accounts for facial affect perception deficits in schizophrenia. This conclusion implies at most a weak link between the perceptual difficulties encountered by people with schizophrenia (e.g., sensory distortions and hallucinations) and facial affect perception difficulties.

## Concluding Comments

Overall, fit to the data is good. The lack-of-fit index, "$\chi^2$,"was validated in that its results were consistent with the results of both $\chi^2$ and *model-mean-*$\chi^2$. In most cases, the best-fitting models appeared to be valid. The exception was *gain*, the parameter associated with the neurotransmitter dysfunction hypothesis. In that case, the model tended to overestimate the variances.

Overall, modeling of latency seemed successful in spite of great variability in the human data. The results demonstrate that *h*, as calculated by the closed-form formula, is a useful parameter. The proportion of error accounted for by latency is only somewhat higher than that accounted for by response content. For example, for the control group (unperturbed networks), the proportions of "$\chi^2$" contributed by the latency component were 62.2 per cent for the first run and 57.3 per cent for the second. This finding is not surprising given that latency data tend to be inherently more variable than content data. Moreover, the backpropagation algorithm explicitly trained the network on (cross-validated) content. The model still provided an adequate fit for the latencies by the estimation parameter ("$\chi^2_{35}$" = 42.64 and 38.57, both *NS*), but $rSS_{latency}$ tended to be greater than one (indicating a poor fit), while $rSS_{accuracy}$ tended to be less than one (indicating a good fit).

In many cases, parameter values seemed to be identifiable (in the mathematical sense, within this study and as estimated with the "$\chi^2$" lack-of-fit criterion). For estimated parameters, the algorithm typically found the same, specific values across different runs. Similarly, the value for *h* was remarkably consistent. The exception was for the control group on the second run. Note, however, that the recalculated values for *h* for the patient groups on the second run were around 0.32, similar to the control group on the first run.

Pages 86 - 89

## CHAPTER FIVE: STUDIES WITH PICTURES OF HUMAN FACES

Results from the previous study are consistent with previous research that suggested that damage to a connectionist network can account for schizophrenic responding. Further, a model accounting for both response latency and response content was found to be acceptable for both patient and nonpatient data by an orthodox sufficiency test. The major change in the next phase of the project was that the stimuli were pictures of realistic human faces instead of line-drawn schematic faces. Two tasks had to be completed before modeling could begin. The first was to describe each stimulus in terms of the probabilities that participants perceived specific features. The next task was to collect data regarding actual judgments – in terms of both latency and accuracy – from patient and nonpatient populations. Moreover, the results from these studies are relevant to schizophrenic facial affect perception in their own right.

Two complications arise from the decision to use pictures of human faces as stimuli. First, the pictures to be studied need to be selected. Much previous research used posed photographs or slides.[9] The current project focuses on four prototypical expressions that tend not to be confused: happy, surprised, sad, and disgusted.

The second difficulty is harder to resolve. The selected stimuli had to be transformed into a format that the computer simulation could access while maintaining psychological validity. (See Vickers & M.D. Lee, 2000, for a discussion of the representation problem in connectionist networks.) This goal was accomplished by identifying the features present in each image. The process involved a qualitative analysis of the range of possible features, and a quantitative calculation of the likelihood of participants perceiving each feature in each image.

### Identifying Features

#### Method

The first task of identifying features proceeded in two phases. First, participants generated a list of possible features. Second, another group of participants marked which of these features were present in each image. An additional study (see Appendix B) resulted in a description of each image in terms of a short list of unambiguous features. Appendix C contains samples of all testing materials.

---

[9] e.g., Archer, Hay, & Young, 1992; Borod, et al., 1989; T.E. Feinberg et al., 1986; Flack, Cavallaro, Laird, & Miller, 1997; Gaebel & Wölwer, 1992; Heimberg, Gur, Erwin, Shtasel, & Gur, 1992; Kee et al., 1998; Kerr & Neale, 1993; Kline et al., 1992; Lewis & Garver, 1994; Mandal & Rai, 1987; Novic, Luchins, & Perline, 1984; Pilowsky & Bassett, 1980; E. Walker et al., 1980; Zuroff & Colussy, 1986.

### Participants

All participants were drawn from the Introductory Psychology Subject Pool at the University of Western Ontario, and each received one "research credit" for participation. Twenty-three students (9 male, 14 female) with a mean age of 20.0 years (standard deviation of 1.61 years) completed phase one. Twenty-five students (10 male, 15 female) with a mean age of 20.5 years (standard deviation of 3.45 years) completed phase two. Students participated in groups, ranging in size from one to five people. All participants had normal visual acuity or lenses that corrected to normal acuity.

### Procedure

After signing an informed consent form, participants were given specific instructions, dependent on the phase. Phase one participants were given examples of features drawn from the literature (such as, eyelids closing, mouth open, nose wrinkled, eyebrows curved and high, and eyebrows raised and flattened) and then asked to list as many features as possible for each image. Phase two participants were given time to examine the list of features and then asked to check off as many features as were present for each image. Stimuli from Ekman and Friesen (1976; see Table C1) were presented as black-and-white slides.[10] Each slide was left until all participants were apparently finished (on average, about one and a half minutes for phase one and three minutes for phase two). All participants were shown the full set of twenty images in the same random order, except the starting point varied systematically.

## Results

Phase one participants generated a substantial list of features. Features taken from the literature were considered as well. For example, five features are known to stimulate specific cells among macaque monkeys: head lowering and eyelids closing (Rolls, 1992), mouth open and grimacing (Perrett et al., 1988), and raised eyebrows (Perrett & Mistlin, 1990). In addition, Ekman and Friesen (1975) cited wide-open eyes, dropped jaw, eyebrows curved and high, eyebrows raised and flattened, corners of lips drawn back and up, lips parted, Crow's-feet wrinkling around the eyes, upper lip raised, wrinkled nose, lower lip raised and slightly forward, lower lip lowered and slightly forward, and eyebrows together as relevant features. Qualitative analysis of phase one data resulted in a list of 125 features. Appendix D provides the data, which consist of the list of features (Table D1), and the proportion of participants who indicated that a given feature was present in a given image (Table D2).

---

[10] My thanks to John Cesarini, who converted the photographs to slides.

# CHAPTER SEVEN: STUDIES OF OTHER TASKS

A critical questions remains: Are the schizophrenic deficits specific to facial affect perception, or are they merely one aspect of a general deficit? Two tasks that are common in schizophrenia research were examined to address this question. The WCST is often used as a test of frontal lobe dysfunction, including that found in schizophrenia. Brain damage has been modeled in connectionist networks by eliminating units (e.g., Hinton & Shallice, 1991). The WCST has been previously modeled with ART networks (see Parks & Levine, 1998). As well, Dehaene and Changeux (1991, cited in Parks & Levine, 1998) developed a neural network model of the WCST based on neuropsychological considerations. CPT is a test of attention, and certain variants seem to be able to identify people with schizophrenia. It has been simulated in a backpropagation network by adjusting *gain* (Cohen & Servan-Schreiber, 1992). In the final two studies, simulations of the WCST and CPT were created and damaged.

These simulations should be considered exploratory. On the one hand, they provide further insight into the effects of the types of damage that had been applied to models of facial affect perception tasks. They also demonstrate the potential for applying formal mathematical modeling criteria to a variety of connectionist models. On the other hand, pragmatic considerations weighed heavily in determining the configuration of the networks and training procedures. In contrast, the application of damage was theoretically driven. Therefore, conclusions will be drawn regarding the effects of damage, but not with respect to these particular instantiations of the tasks themselves. The use of published data precluded cross-validation techniques.

## Wisconsin Card Sorting Task

This analysis is based on Kimura's (1986) instructions for the administration of the WCST. The patient is instructed to put a deck of 128 cards into four piles, corresponding to each of four stimulus cards. Every card has one of four symbols (triangle, star, cross, or circle) printed in one of four colors (red, green, yellow, or blue), repeated one to four times. All possible combinations are repeated twice. The stimulus cards are one red triangle, two green stars, three yellow crosses, and four blue circles. After the patient has placed each card, the tester responds *Right* or *Wrong*. No other feedback or instruction is provided. The first sorting category is color. After ten consecutive cards have been correctly sorted by color, the sorting category shifts to form. The patient is not informed of the shift, except by listening to the tester's responses. After ten consecutive cards have been sorted by form, the sorting category shifts to number, followed by color, and then form, number, and color again. The test continues until the patient correctly sorts through all six sorting categories or until the patient attempts all 128 cards.

The tester records all responses, indicating the nature of the sort (color, form, number, some combination, or unique) and whether it was correct. Responses are scored as one of six types. The first error in a new category is an *information error*. For example, after someone correctly sorts ten cards by color and then categorizes the next card the same way, that eleventh trial is an information error. The card should be sorted by form, but the person has no way of knowing that the sorting dimension has changed until after placing the eleventh card. Only the ten consecutive correct responses are scored as *correct responses*. Other correct responses are called *inadvertent* or *surplus correct responses*. Errors can be classified as perseverative or nonperseverative. Incorrect responses that would have been correct in the immediately preceding category are *perseverative errors*, except for the first, which is an information error. Responses which are not correct and not perseverative, including unique responses, are *nonperseverative errors*.

Sullivan et al. (1993) examined the structure of the WCST. They found that the responses of people with schizophrenia (*n* = 22) and people with frontal lobe damage (from a previous study) tended to load on a Perseveration factor. The responses of people who experienced chronic alcoholism (*n* = 20) tended to load on an Ineffective Sorting factor. Members of the control group (*n* = 16) made significantly fewer perseverative errors than the schizophrenia group, made significantly fewer inadvertent/surplus correct responses than the alcoholism group, and achieved significantly more categories than both the schizophrenia and alcoholism groups. Table 10 shows the results of their study that are relevant to the current simulation. Note that Sullivan et al. administered the WCST atypically in that their subjects sorted all 128 cards.

Intuitively, instantiating the WCST in a backpropagation network would seem to require some “memory” of previous trials and some feedback regarding the previous response. Elman (1990) solved the problem of memory in backpropagation networks by adding context units. In an Elman network, the number of context units equals the number of hidden units. The activation of each context unit is set at the previous activation of the corresponding hidden unit. Similarly, information regarding the previous response can be provided simply by connections leading from the output units to the hidden units (see Jordan, 1992; Jordan & Rumelhart, 1992). Jordan networks have been particularly successful in simulating motor tasks.

Table 10

Wisconsin Card Sorting Task data[a]

| | Nonpatient Controls | Schizophrenic Patients |
|---|---|---|
| Number of Categories Achieved[b] | 8.3 | 4.6 |
| Correct Responses[c] | 83[d] | 46 |
| Inadvertent / Surplus Correct | 18 | 30 |
| Perseverative Errors | 8 | 29 |
| Nonperseverative Errors | 19 | 23 |

Notes. [a] Data based on Sullivan et al., 1993.

[b] Number of Categories Achieved is provided for information only, and does not participate in the model fit.

[c] Correct Responses excludes Inadvertent / Surplus Correct responses.

[d] Values are rounded off to integer values and adjusted to sum to exactly 128.

In the current project, the first attempts to simulate the WCST were networks based on the architectures developed by Elman and by Jordan. A wide range of variations of these architectures were tried, but each failed. As with the unsuccessful model discussed at the end of Chapter Six, these failures do not necessarily imply that there is no Elman- or Jordan- based architecture that could simulate the WCST, or that the architectures attempted would not succeed with further or different training. Empirically, however, these architectures did not seem well suited to the WCST. The difficulty seemed to be related to the problem of catastrophic interference (see, e.g., Long, Parks, & Levine, 1998, pp. 23 - 25). Learning later dimensions may have interfered with previously learned dimensions. Additional pretraining in which the sorting dimension was changed every few epochs (instead of within each epoch) did not alleviate this situation.

To facilitate the model, context was provided somewhat artificially (see below). Obviously, this solution is less than ideal. It is, however, acceptable for the task at hand for the following reasons. Most important, the purpose of the current investigation is to examine what modifications to a network that simulates nonpatient data lead to patient-like responding. Further, the mechanism for providing context is equally artificial for both groups. As previously discussed, the backpropagation algorithm is rather artificial anyway. Most notable in this case is the need for multiple training epochs: Human data is based on the first attempt of the task. (See Vickers & M.D. Lee, 2000, for a discussion of this learning problem.)

## Method

### Architecture

The model consisted of four layers, as depicted in Figure 15. The 12 units in the input layer represented the stimulus card. They sent activations to the ten units in the hidden layer and to the four context units. The four context units also sent activations to the hidden layer. The hidden layer sent activations to the four output units.

### Representation

Each stimulus card was represented by a unique pattern of activations in the input layer. The twelve units were divided into three groups of four. The three groups represented the three dimensions (i.e., color, form, and number), and each unit represented one element within that dimension (e.g., red). Each output unit represented one of the four possible discard piles.

### Learning Rule and Associated Parameters

The WCST was simulated with a backpropagation network (using Carter, 1997 - 1999). For pretraining, $\varepsilon$ was set to .1, and the momentum for changes in weights was set to .4. For subsequent training, $\varepsilon$ was reduced to .01, and momentum for changes in weights was set to zero for one trial, but then returned to .4.

## Processing of a Pattern

The activations of input units were set for each trial. Hidden and output unit activations were set to zero at the beginning of each epoch. The activations of the four context units were determined by the simulation, as follows. One context unit was active if and only if the previous response had been correct. The other three each represented a dimension, and they indicated the sorting criterion that had apparently been used on the previous trial. When only one dimension from the input had the same active element as the output layer, then the context unit representing that dimension was active for the next trial. For each iteration, all layers were updated sequentially, beginning with input, followed by context and hidden, and ending with output.

If more than one dimension had the same active element as the output layer then the following rules applied. If there had been an information error and the correct sorting dimension was eligible, then the correct dimension unit was active for the next trial. (For example, if the simulation had incorrectly sorted the previous card on form, having previously made ten correct responses on form, and on the current trial the third units of both the number and form input unit sets were active and the third output unit had the highest activation, then the response would be scored as *Correct* and the Number context unit would be active for the next trial.) If there had not been an information error and the perseverative dimension was eligible, then the perseverative dimension unit was active for the next trial. (Continuing the example, if the simulation went on to correctly sort nine more cards on number, and on the current trial the third units of the number and color input unit sets were both active and the third output unit had the highest activation, then the response would be scored as *Incorrect*, and the Number context unit would still be active for the next trial.) In all other cases, no dimension unit was active for the next trial.

The interpretation of the network's output involved a comparison of the output units' relative activations. Each output unit's activation was viewed as an independent probability estimate that the category represented by that unit is correct[11] (instead of the more common interpretation of the activation of all the output units as a pattern). Selection of a category by the network was inferred by a simple decision rule: The output unit that accrues the highest activation was declared the network's choice.

## Training Regime

McRae and Hetherington (1993) demonstrated that pretraining a network can have a profound effect on subsequent learning. A pilot test of the WCST simulation suggested that pretraining would be advantageous because the network required a very large number of training epochs to learn the first criterion. One potential reason for this difficulty is that the naïve network does not know, for example, to associate a red stimulus with the category Red. Human

---

[11] My thanks to David Vollick for suggesting this interpretation. See also Elman (1990, pp. 197 - 198).

subjects, of course, readily make such associations. Therefore, the network was trained to associate one input unit from each of the three sets of input units with each of the output units. This task was accomplished by setting the activation of one input unit at a time to one, with the activations for the remaining input units and all context units set to zero. The target activations of the output units were one for the unit corresponding to the correct category and zero for the other units. The network was trained until it made no errors on this preliminary task. Weights were updated at the end of each epoch.

Unidimensional data, with only one input unit active for each trial, were presented in random order. By epoch 1, 599, the backpropagation error was 6.235 and the simulation consistently made no errors. At that point, training continued with multidimensional input. Trials were still in random order, and weights were still updated at the end of each epoch. After each sequence of ten correct responses, the sorting dimension was switched to a different dimension at random. At epoch 282, 076, backpropation error was 28.6212, and $\chi^2_3$ was 1.08 (*NS*). Training continued, but after each sequence of ten correct responses, the training dimension switched to the next dimension in order (color, form, number, etc.).

The feedback patterns were atypical. The template was different for correct and incorrect trials. Human subjects are told if their responses are *Correct* or *Wrong*, but not the correct discard pile. To simulate this feedback, on correct trials the network was given feedback with the correct output unit as one and the other three as zero. On incorrect trials, however, the network was given merely the information that the chosen output unit was incorrect (zero), and one of the other output units was correct, but not which one (all the same value, one). Taken literally, this feedback would imply that the network should select all three of the other categories. Over a number of trials, however, the activation of the output units converged on the mean feedback.

## Basic Analyses

Error was calculated according to the following legitimate $\chi^2$ formula:

$$\chi^2_{df=3} = \sum_{i=1}^{4} \frac{(N_{i,model} - N_{i,humans})^2}{N_{i,model}}$$

where $N_{i,\,model}$ is the number of trials categorized as type *i* (i.e., correct responses, surplus correct responses, perseverative errors, and nonperseverative errors) made by the *model*, with the second subscript changed for the *human* participants. Human data were based on Sullivan et al. (1993). The estimation procedure for damage parameters was the same as for previous simulations. Unlike previous simulations, there is only one set of data, making cross-validation procedures impossible. No degrees of freedom were lost in estimating the backpropagation parameters, however, because the network was trained on ideal responding rather than the actual human data. Because this formula does not depend upon the model variance, testing *z* for artificially fitting the data is not necessary. Further, this study avoids the complexities of interpretation imposed by alternate lack-of-fit measures. This $\chi^2$ formula corrects for scale, making it superior to *SS*. Unlike $\hat{\chi}^2$, the model values in the denominator are constrained: They must sum to the total number of trials. Therefore, any advantage gained by overestimates in one category are penalized by underestimates in all of the others. Lack-of-fit and change-in-fit probabilities can be calculated and interpreted with the usual level of confidence because this $\chi^2$ is legitimate. Of course, other lack-of-fit criteria, such as squared correlations or *rSS* could have been applied, but were expected to yield little new information relative to the effort required to implement them at this time.

## Results

The simulation was able adequately to replicate the data collected from the control group. The best fit was achieved at epoch 712, 216, when backpropagation error was 32.6 and $\chi^2_3$ was 1.92 (*NS*). This value is less than the median value for the $\chi^2_3$ distribution (2.37, i.e., $p > .5$). The model achieved 80 correct responses, 21 surplus correct responses, 11 perseverative errors, and 16 nonperseverative errors.

The unperturbed model was not suitable for the data collected from people with schizophrenia ($\chi^2_3 = 41.96, p < .001$). Cutting 19 per cent of the weights, specifically the putatively least important (lowest squared value, the lowest $\partial E / \partial w$, or the lowest $[\partial E / \partial w]^2$), provided the most improvement in fit and led to model results that were not statistically-significantly different from the observed data ($\chi^2_2 = 3.19$, *NS*; $\Delta\chi^2_1 = 38.77, p < .001$). Similarly, cutting 9 per cent of the “least important” units, specifically those with the lowest sum of squared efferent weights, the lowest $\partial E / \partial a$, or the lowest $(\partial E / \partial a)^2$ also improved the fit and yielded results similar to the observed data ($\chi^2_2 = 4.52$, *NS*; $\Delta\chi^2_1 = 37.44, p < .001$; in all cases, the output units were protected from being eliminated).

Several other types of damage also improved the fit, but the resulting models remained statistically-significantly different from the observed data. Adjusting *gain* to 0.81 and multiplying all weights by 0.82 both led to the same level of improvement ($\chi^2_2 = 7.62, p < .025$; $\Delta\chi^2_1 = 34.34, p < .001$). Adding noise to the hidden units at each tick was also partially effective, particularly from a rectangular distribution with a standard deviation of 0.091 ($\chi^2_2 =$ 11.8 [S.D. = 8.4844], $p < .005$; $\Delta\chi^2_1 = 30.16, p < .001$). This finding was replicated with a Gaussian distribution and with a rectangular distribution applied to the output units.

## Continuous Performance Task

The Continuous Performance Test (CPT; Rosvold et al., 1956) was developed to identify brain damage (broadly defined). Although physiological evidence suggested that patients with brain damage should have an inferior ability to sustain attention, their performance was inconsistent on relevant intelligence subtests (such as Digit Span and Digit Symbol). In the CPT, the participant is shown each member of a series of symbols for a set time interval. The original two variants are CPT-X and CPT-AX. For CPT-X, the participant presses the response key if and only if the letter *X* was presented. CPT-AX is inherently more difficult: The participant presses the response key if and only if the letter *X* follows a letter *A*. A more recent variant, CPT-IP (Identical Pairs; Cornblatt, et al., 1988, also known as *CPT-sequential target version*, Friedman, Vaughan, & Erlenmeyer-Kimling, 1981), requires that the participant "respond to any digit 0 through 9 that repeated on successive trials" (Strandburg et al., 1994, p. 527). The proportion of target trials for CPT-IP is usually .2. CPT-IP yields several different descriptive statistics. For the present purposes, the most informative are *hits* (i.e., the proportion of target trials to which the participant responded) and *false alarms* (i.e., the proportion of nontarget trials to which the participant responded).

Rosvold et al., (1956) alluded to unpublished data that indicated that people with schizophrenia performed about as well as people without the disorder on CPT-X and CPT-AX. CPT-IP, however, does identify people who have schizophrenia, whether they are adults (Cornblatt, Lazenweger, & Erlenmeyer-Kimling, 1989) or children (Strandburg et al., 1994). Indeed, children with schizophrenia apparently respond similarly to their adult counterparts on CPT tasks (Strandburg et al.). Strandburg et al. found that children with schizophrenia had a smaller proportion of hits than did nonpatients (.73 versus .89), and a greater proportion of false alarms (.0285 versus .00575) over an epoch of 400 trials.

## Method

### Architecture

Several network configurations and training procedures were attempted. The successful CPT-IP network, depicted in Figure 16, had seven input units, two layers of hidden units, and a single output unit. The first hidden layer consisted of six units. It received activation from input units and from the second hidden layer. The first hidden layer sent activation to the output unit. The second hidden layer consisted of four units. It received activation from the input unit and sent activation to the first hidden layer.

### Representations

Stimuli analogous to those used by Strandburg et al. (1994) were generated by considering the seven positions common to simple digital displays. Digits from *0* to *9* (excluding *8*) were generated by activating the units corresponding to the display, as shown in Figure 17. Each of these simple and distinct patterns consists of seven bits, each representing a bar from the visual display. In this way, the visual similarity of the digits is retained. The digit *8* is excluded: Strandburg et al. excluded *8* because it had served as the target stimulus in a CPT-X task with the same participants. Activation of the output unit represented a response by the network.

### Learning Rule and Associated Parameters

CPT-IP was simulated with a backpropagation network (using Carter, 1997 - 1999). The learning rate, ε, was .1 throughout training. Backpropagation error fell from 51.5 at epoch 1 to 26.34 at epoch 52, 000. At that point, $\chi^2_2$ was 89.1, and training seemed to have stalled. The momentum for changes in weights was .4 up to epoch 52, 000 and zero thereafter.

## Processing of a Pattern

The activations of input units were set for each trial. Hidden and output unit activations were set to zero at the beginning of each epoch. The simulation was interpreted to have made a response when the output unit's activation was greater than a cutting score, Y, that was initially set, arbitrarily, to .5. The cutting score was estimated periodically after epoch 52, 000. At some points in training, adjusting Y improved fit, but not enough for an adequate fit overall. In the end, the default value for Y was sufficient, and other values did not lead to any improvement in fit. For each iteration, all layers were updated sequentially, beginning with the input layer, followed by the first and then second hidden layers, and ending with the output layer.

## Training Regime

Each epoch had 400 trials. Initially, and for most of the training phase, the probability that a given trial was a target (i.e., the stimulus was the same as in the previous trial) was .5 to ensure that there were approximately equal numbers of target and nontarget trials. Up to epoch 52, 000, the same series of stimuli was used. Thereafter, a new series of stimuli was generated periodically. At 105, 620 epochs, backpropagation error was 9.34, and the probability of a trial being a target was set to .2. Weights were updated at the end of each epoch. The training value for the output unit was one if the current trial was a target and zero otherwise.

## Basic Analyses

Error was calculated as a legitimate $\chi^2$ according to the following formula:

$$\chi^2_{df=2}=\sum_{i=1}^{2} N_i[\frac{(p_{correct,model,i}-p_{correct,human,i})^2}{p_{correct,model,i}}+\frac{(p_{error,model,i}-p_{error,human,i})^2}{p_{error,model,i}}]$$

where $N_i$ is the number of trials in category $i = 1$ (target) or $i = 2$ (not target), and $p_{correct,\ model,\ i}$ is the probability of the model making a correct response on trials of that category (with subscripts changed for *human* and *error* trials). Human data was taken from Strandburg et al. (1994). As with the WCST simulation, no degrees of freedom were lost when estimating the backpropagation parameters because the network was trained on ideal, rather than human, responding. This $\chi^2$ is legitimate, allowing for inferences regarding lack-of-fit and change-in-fit to be drawn with the usual level of confidence.

## Results

The simulation was able adequately to replicate the nonpatient data. The best fit was achieved at epoch 106, 938, when backpropagation error was 7.99, and $\chi^2_2$ was 1.19 (*NS*). This value is less than the median value for the $\chi^2_2$ distribution (which is 1.39; therefore, $p > .5$). At that point, the simulation responded to .8875 of the target trials and .0125 of the nontarget trials, as compared to .8875 and .00575, respectively, for the children in the control group.

The unperturbed model was not appropriate for the data collected from people with schizophrenia ($\chi^2_2 = 26.51, p < .001$). Because of the limited degrees of freedom, each damage parameter was estimated individually (in contrast to previous simulations where, for example, means and standard deviations of noise distributions were estimated concurrently).

The best fit was achieved by setting *gain* to 0.7308. The model responded to .775 of the target trials and .0219 of the nontarget trials, whereas the patients responded to .73 of the target trials and .0285 of the nontarget trials. The difference is not statistically significant ($\chi^2_1 = 1.59$, *NS*), and it represents a statistically significant improvement over the undamaged model ($\Delta\chi^2_1 = 24.92, p < .001$).

Several other types of damage resulted in smaller improvements in fit. Not one was as effective as adjusting *gain*, and not one resulted in an adequate fit to the data obtained from the patients with schizophrenia. Deleting the least important weights, either the .009 with the lowest squared value or the .091 with either the smallest $\partial E / \partial w$ or the smallest $(\partial E / \partial w)^2$, improved the fit (in all three cases, overall fit $\chi^2_1 = 7.57, p < .01$, and improvement $\Delta\chi^2_1 = 18.94, p < .001$). Adding a constant, 0.001, to the input or hidden layer activations yielded the same improvement. Distorting the weights by multiplying them by a number drawn from a rectangular distribution with a mean of one and standard deviation of 0.004 gave $\chi^2_1 = 17.67$ (S.D. = 8.9112; $p < .005$; $\Delta\chi^2_1 = 8.84, p < .005$). The result with a Gaussian distribution was similar. A somewhat better fit was found by distorting the weights by multiplying them by .667, yielding $\chi^2_1 = 5.7$ ($p < .025$; $\Delta\chi^2_1 = 20.81, p < .001$). Adding noise from a rectangular distribution and a standard deviation of 0.0624 to the hidden units also improved the fit ($\chi^2_1 = 12.29$ [S.D. = 4.9892], $p < .001$; $\Delta\chi^2_1 = 14.22, p < .001$), but a Gaussian distribution was slightly better. An even better fit was achieved by adding noise from a rectangular distribution with a standard deviation of 0.0625 to the output unit ($\chi^2_1 = 3.7$ [S.D. = 1.9851], $.1 > p > .05$; $\Delta\chi^2_1 = 22.81, p < .001$). A Gaussian distribution did almost as well. Neither cutting units nor adding noise to the input units improved the fit.

## Discussion

These simulations of the WCST and the CPT-IP illustrate that apparently similar deficits may derive from different processes. These results highlight the importance of attempting different interventions within each simulation and of testing different tasks with the same set of interventions.

One strength of these simulations is their simplicity. Indeed, the WCST network is considerably simpler than those described by Parks and Levine (1998; notwithstanding the atypical algorithm used here to determine the activations of the context units). Similarly, the CPT-IP simulation is also quite straightforward. More complex architectures were attempted in each case, but they failed to reproduce the observed data adequately.

Adjusting *gain* improved the fit in both tasks. It was particularly beneficial in replicating deficits found among people with schizophrenia on the CPT-IP. Indeed, it was the only type of damage that led to model responses that were not statistically- significantly different from the observed data for the CPT-IP. The WCST simulation partially replicated the finding by Leven and Levine (1995, cited in Parks & Levine, 1998) in that reducing *gain* improved the fit towards matching the deficits that were found in the clinical group (patients with schizophrenia here and patients with frontal lobe damage in their study). The data from the WCST simulation with reduced *gain*, however, was still statistically-significantly different from the data observed among the people with schizophrenia.

In contrast, the most effective damage for the WCST simulation was to eliminate certain components of the network. Specifically, removing the least important weights (as measured by either their squared value or their influence on error) led to an acceptable fit. Removing the units with the lowest sum of squared efferent weights or lowest influence on error was almost as effective. These similar results are not surprising because removing the least effective units effectively removes their efferent weights. In contrast, cutting units was not effective at all for the CPT-IP simulation.

These findings suggest that the deficits measured by the CPT-IP and WCST – although appearing to be similar – may derive from different physiological mechanisms. The most effective parameter for the CPT-IP was *gain*, which may correspond to neurochemical anomalies. In contrast, the most effective parameters for the WCST were related to cutting weights and units, which may correspond to structural damage. The damage to the WCST simulation was systematic rather than random. This finding provides support for the neural Darwinist perspective, and tends to oppose the idea that damage is arbitrarily determined by some purely physiological disease process. In particular, the weights or units removed were the least important (by any one of several criteria). This finding suggests that the normal process of synaptic pruning simply goes too far in people with schizophrenia.

# APPENDIX D: NORMATIVE DATA ON FEATURES

Table D1

List of features

1. Head lowering
2. Head tilted to one side
3. Furrowed brow
4. Forehead wrinkled
5. Eyebrows lowered
6. Eyebrows lowered and straight
7. Eyebrows lowered and pointing down to the outside
8. Eyebrows low and arched
9. Eyebrows low and far apart
10. Eyebrows level and straight
11. Eyebrows raised
12. Eyebrows raised and flattened
13. Eyebrows curved and high
14. Eyebrows together
15. Eyebrows slanted slightly upwards
16. Eyebrows slanted inward
17. Inside of eyebrows slant down slightly
18. One eyebrow crooked
19. Open eyes
20. Wide-open eyes
21. “Bug-eyed”
22. Eyes very small
23. Eyelids closing
24. Eyes slightly closed
25. Eyes three-quarters open
26. Eyes half open
27. Eyes almost squinted shut
28. Eyes slightly squinting
29. One eye slightly more open
30. Lower eyelid raised
31. Corners of eyes drawn down
32. Eyes happy and vibrant
33. Eyes small and slit-like
34. Sparkle in the eyes
35. Scary-looking eyes
36. Sunken eyes
37. Glaring
38. Eyes droopy

39. Looking downward
40. Irus low in eye
41. Irus high in eye
42. Staring transfixed
43. Eyes focused straight ahead
44. Mouth open
45. Mouth wide open
46. Mouth slack and open
47. Grimacing
48. Pouting
49. Frowning
50. Snarling
51. Sneering
52. Slight frown
53. Slight smile
54. Wide smile
55. Mouth straight
56. Mouth closed and pulled up at corners
57. Mouth closed and turned down at corners
58. Mouth looks like its growling
59. Lips parted
60. Lips tight together
61. Lips together
62. Lips together and pushed up
63. Lips together pushed down
64. Lips pursed
65. Lips soft
66. Middle part of the upper lip is downward
67. Slight tension at mouth corners
68. Corners of lips drawn back and up
69. Corners of mouth pointing down
70. Corners of mouth pulled up tightly
71. Lower lip protruding slightly
72. Lower lip pulled up
73. Outside corners of the mouth go way up
74. Mouth pulling up & to one side
75. Wrinkled upper lip
76. Upper lip slightly raised
77. Upper lip raised but flat across
78. Upper lip pulled up exposing front teeth entirely
79. Teeth showing
80. Teeth and gums showing
81. Lower teeth showing slightly
82. Lower teeth showing

83. Upper teeth showing
84. Teeth together
85. Tongue showing
86. Tongue drawn back
87. Tongue sticking behind lips
88. Muscles bulging under bottom lip
89. Bulging chin muscles
90. Nose wrinkled
91. Wrinkles beside nose
92. Nose pinched
93. Nose snarling
94. Nose curled
95. Nose drawn up and wrinkled
96. Lines below the nose on one side
97. Wrinkles around mouth
98. Wrinkle between cheeks and nose
99. Wrinkle in jawline and cheeks
100. Lines on one side of the mouth
101. Lines from nose and chin make a diamond shape
102. Deep line around nose going down to the chin area
103. Skin under eyes wrinkled
104. Wrinkles showing on the inside corner of the eyes
105. Crow's-feet wrinkling around the eyes
106. Wrinkle between eyes and above nose
107. Muscles around eyes contracted
108. Bags under the eyes
109. Wrinkled temples
110. Cheeks raised
111. Cheeks pulled back
112. Cheeks sunken in
113. Cheeks bulging out
114. Dimple showing
115. Nostrils flared
116. Nose tip pointed down
117. Double chin showing
118. Chin bunched up
119. Chin protruding
120. Chin drawn in
121. Dropped jaw
122. Clenched jaw
123. All features directed towards middle of face
124. Tension in the face
125. Bottom quarter of face flattened and stretched horizontally

Table D2

Probability that a given feature (columns) was endorsed for a given face (rows), collapsed across subjects.

| Face | Features 001 to 015 | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 01 | .00 | .00 | .60 | .40 | .48 | .20 | .04 | .04 | .08 | .12 | .00 | .00 | .00 | .08 | .20 |
| 02 | .00 | .04 | .12 | .00 | .08 | .00 | .08 | .16 | .12 | .36 | .00 | .04 | .00 | .04 | .12 |
| 03 | .00 | .28 | .12 | .08 | .08 | .00 | .08 | .28 | .08 | .00 | .16 | .04 | .08 | .00 | .12 |
| 04 | .04 | .12 | .04 | .00 | .32 | .04 | .00 | .36 | .08 | .08 | .04 | .00 | .00 | .00 | .00 |
| 05 | .00 | .04 | .08 | .72 | .00 | .00 | .00 | .00 | .00 | .00 | .68 | .16 | .52 | .00 | .08 |
| 06 | .04 | .76 | .28 | .16 | .24 | .04 | .32 | .08 | .08 | .00 | .04 | .00 | .00 | .04 | .52 |
| 07 | .00 | .04 | .08 | .76 | .00 | .00 | .00 | .00 | .08 | .00 | .60 | .08 | .64 | .00 | .04 |
| 08 | .04 | .00 | .20 | .08 | .48 | .00 | .12 | .32 | .04 | .00 | .00 | .00 | .00 | .04 | .00 |
| 09 | .04 | .00 | .12 | .44 | .52 | .48 | .12 | .04 | .04 | .24 | .00 | .00 | .00 | .04 | .00 |
| 10 | .00 | .00 | .16 | .08 | .64 | .08 | .12 | .20 | .16 | .04 | .00 | .00 | .00 | .04 | .04 |
| 11 | .00 | .00 | .04 | .84 | .00 | .08 | .04 | .04 | .00 | .00 | .60 | .32 | .20 | .00 | .04 |
| 12 | .00 | .00 | .08 | .00 | .12 | .00 | .04 | .24 | .08 | .00 | .20 | .00 | .24 | .00 | .28 |
| 13 | .24 | .00 | .32 | .04 | .52 | .04 | .24 | .32 | .08 | .04 | .00 | .00 | .00 | .00 | .12 |
| 14 | .00 | .32 | .04 | .00 | .40 | .48 | .16 | .16 | .00 | .08 | .00 | .00 | .00 | .00 | .00 |
| 15 | .00 | .12 | .04 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .48 | .00 | .72 | .00 | .00 |
| 16 | .12 | .00 | .16 | .12 | .48 | .60 | .04 | .00 | .08 | .04 | .00 | .00 | .00 | .00 | .00 |
| 17 | .00 | .08 | .16 | .04 | .36 | .40 | .04 | .04 | .00 | .36 | .00 | .00 | .00 | .00 | .00 |
| 18 | .00 | .00 | .08 | .56 | .00 | .00 | .00 | .00 | .00 | .00 | .52 | .04 | .72 | .00 | .04 |
| 19 | .00 | .04 | .12 | .72 | .00 | .00 | .00 | .00 | .00 | .00 | .56 | .00 | .48 | .00 | .16 |
| 20 | .00 | .04 | .04 | .00 | .28 | .44 | .16 | .08 | .12 | .16 | .00 | .04 | .00 | .00 | .00 |

| Face | Features 016 to 030 |
|---|---|
| 01 | .12 .16 .08 .20 .00 .00 .00 .20 .20 .28 .24 .00 .16 .00 .08 |
| 02 | .00 .08 .00 .76 .04 .00 .04 .04 .04 .20 .00 .00 .12 .12 .00 |
| 03 | .00 .00 .08 .40 .08 .00 .08 .00 .08 .04 .20 .12 .32 .04 .04 |
| 04 | .32 .20 .00 .60 .00 .00 .04 .16 .12 .16 .16 .00 .04 .00 .00 |
| 05 | .00 .04 .08 .40 .68 .64 .00 .00 .00 .00 .00 .00 .00 .08 .04 |
| 06 | .04 .00 .00 .80 .00 .04 .00 .04 .00 .16 .12 .00 .04 .00 .00 |
| 07 | .04 .08 .00 .48 .76 .44 .00 .00 .00 .00 .00 .00 .00 .00 .08 |
| 08 | .52 .32 .00 .24 .00 .04 .00 .00 .16 .32 .12 .00 .28 .04 .08 |
| 09 | .08 .12 .04 .28 .00 .04 .16 .08 .08 .24 .20 .00 .16 .20 .00 |
| 10 | .24 .28 .08 .44 .00 .00 .00 .04 .08 .16 .24 .00 .24 .00 .04 |
| 11 | .00 .04 .04 .36 .72 .48 .00 .00 .00 .00 .00 .00 .00 .00 .00 |
| 12 | .08 .16 .20 .40 .00 .00 .20 .04 .12 .16 .16 .00 .24 .40 .04 |
| 13 | .00 .12 .24 .84 .04 .08 .00 .00 .08 .20 .00 .00 .00 .00 .00 |
| 14 | .04 .04 .08 .08 .00 .04 .28 .12 .16 .08 .00 .64 .08 .28 .00 |
| 15 | .00 .08 .04 .84 .40 .08 .00 .00 .00 .08 .04 .00 .04 .00 .00 |
| 16 | .04 .08 .00 .36 .00 .04 .04 .00 .16 .20 .12 .04 .20 .08 .00 |
| 17 | .04 .04 .24 .36 .00 .00 .08 .00 .20 .36 .04 .00 .16 .24 .00 |
| 18 | .00 .08 .12 .40 .84 .20 .00 .00 .00 .00 .00 .00 .00 .00 .04 |
| 19 | .00 .00 .16 .60 .56 .12 .00 .00 .00 .00 .00 .00 .00 .00 .00 |
| 20 | .04 .04 .08 .12 .00 .00 .16 .16 .20 .04 .08 .36 .20 .08 .12 |

| Face | Features 031 to 045 |
|---|---|
| 01 | .36 .00 .08 .00 .04 .48 .08 .40 .20 .12 .20 .20 .40 .00 .00 |
| 02 | .08 .44 .08 .36 .00 .04 .12 .04 .04 .04 .04 .16 .80 .00 .00 |
| 03 | .12 .96 .04 .48 .00 .00 .00 .00 .00 .00 .04 .12 .68 .52 .28 |
| 04 | .12 .00 .08 .00 .20 .52 .32 .36 .20 .04 .04 .20 .40 .00 .00 |
| 05 | .00 .00 .00 .12 .20 .00 .16 .00 .04 .12 .04 .36 .44 .72 .08 |
| 06 | .24 .00 .04 .04 .04 .68 .12 .44 .00 .04 .12 .28 .40 .00 .00 |
| 07 | .00 .28 .00 .28 .00 .00 .08 .00 .00 .00 .04 .28 .44 .88 .16 |
| 08 | .08 .00 .00 .00 .60 .04 .36 .00 .08 .04 .04 .12 .44 .88 .04 |
| 09 | .08 .00 .08 .00 .08 .44 .08 .16 .28 .16 .00 .20 .44 .60 .00 |
| 10 | .20 .00 .00 .00 .24 .08 .28 .04 .04 .04 .00 .12 .60 .00 .00 |
| 11 | .04 .00 .00 .20 .12 .00 .24 .04 .00 .12 .04 .32 .48 .00 .00 |
| 12 | .32 .88 .00 .20 .00 .04 .04 .04 .00 .00 .04 .08 .52 .88 .16 |
| 13 | .12 .00 .00 .00 .00 .52 .00 .44 .56 .16 .00 .24 .24 .00 .00 |
| 14 | .16 .00 .40 .04 .16 .00 .40 .00 .04 .04 .00 .12 .56 .00 .00 |
| 15 | .00 .88 .00 .24 .00 .00 .00 .00 .00 .00 .04 .08 .72 .64 .60 |
| 16 | .04 .00 .00 .04 .52 .12 .52 .00 .04 .00 .16 .28 .48 .00 .00 |
| 17 | .16 .00 .00 .04 .12 .24 .16 .08 .04 .08 .12 .20 .52 .00 .00 |
| 18 | .00 .16 .00 .08 .00 .04 .08 .00 .04 .00 .12 .32 .52 .60 .08 |
| 19 | .04 .00 .00 .24 .08 .08 .24 .00 .00 .00 .12 .52 .48 .80 .04 |
| 20 | .04 .72 .24 .20 .00 .00 .00 .04 .04 .04 .04 .04 .56 .72 .20 |

| Face | Features 046 to 060 |
|---|---|
| 01 | .00 .00 .08 .28 .00 .04 .40 .00 .00 .36 .08 .48 .04 .00 .12 |
| 02 | .00 .04 .00 .00 .00 .00 .00 .80 .00 .52 .52 .00 .00 .00 .28 |
| 03 | .00 .00 .00 .00 .00 .00 .00 .00 .92 .00 .04 .00 .00 .72 .00 |
| 04 | .00 .16 .24 .36 .08 .24 .36 .00 .00 .16 .08 .68 .12 .00 .16 |
| 05 | .36 .08 .00 .00 .00 .00 .04 .00 .00 .20 .00 .00 .00 .76 .00 |
| 06 | .04 .04 .20 .24 .00 .00 .36 .00 .00 .56 .00 .20 .00 .08 .20 |
| 07 | .16 .00 .00 .00 .00 .00 .00 .36 .00 .24 .00 .00 .00 .84 .00 |
| 08 | .00 .12 .00 .04 .68 .12 .08 .00 .00 .00 .04 .00 .52 .80 .00 |
| 09 | .00 .12 .00 .04 .68 .12 .08 .00 .00 .00 .04 .00 .52 .80 .00 |
| 10 | .00 .28 .12 .32 .28 .20 .12 .00 .00 .36 .12 .24 .16 .00 .60 |
| 11 | .00 .04 .04 .00 .00 .04 .00 .24 .00 .80 .20 .00 .00 .00 .24 |
| 12 | .04 .00 .00 .00 .00 .00 .00 .12 .84 .04 .04 .00 .00 .88 .00 |
| 13 | .00 .08 .40 .36 .00 .00 .36 .00 .00 .04 .04 .68 .00 .00 .16 |
| 14 | .00 .32 .04 .12 .48 .40 .04 .00 .00 .00 .08 .28 .48 .00 .40 |
| 15 | .00 .04 .00 .00 .00 .00 .00 .08 .56 .04 .00 .00 .00 .64 .00 |
| 16 | .00 .28 .04 .28 .52 .20 .12 .00 .00 .08 .00 .60 .36 .00 .76 |
| 17 | .00 .36 .20 .24 .04 .16 .12 .04 .00 .28 .20 .28 .04 .00 .28 |
| 18 | .52 .00 .04 .00 .04 .00 .00 .16 .00 .04 .00 .00 .00 .88 .00 |
| 19 | .36 .04 .00 .00 .00 .00 .08 .00 .00 .36 .00 .00 .04 .92 .00 |
| 20 | .00 .00 .00 .00 .00 .00 .00 .04 .92 .08 .12 .00 .00 .88 .00 |

| Face | Features 061 to 075 |
|---|---|
| 01 | .80 .00 .20 .00 .12 .12 .04 .00 .36 .00 .04 .00 .00 .00 .04 |
| 02 | .64 .12 .00 .04 .20 .12 .00 .08 .00 .12 .00 .00 .00 .08 .00 |
| 03 | .00 .00 .00 .00 .00 .40 .08 .24 .00 .28 .08 .04 .32 .00 .00 |
| 04 | .72 .04 .32 .08 .00 .24 .12 .00 .40 .00 .16 .00 .00 .04 .00 |
| 05 | .00 .00 .00 .00 .12 .32 .08 .08 .08 .00 .00 .04 .00 .00 .00 |
| 06 | .72 .00 .08 .00 .12 .20 .20 .00 .00 .00 .00 .00 .00 .00 .04 |
| 07 | .00 .00 .00 .00 .20 .20 .00 .12 .00 .00 .16 .04 .00 .00 .00 |
| 08 | .00 .00 .00 .04 .08 .16 .16 .04 .12 .00 .12 .08 .04 .00 .24 |
| 09 | .08 .00 .04 .00 .08 .32 .24 .08 .04 .00 .00 .00 .00 .08 .04 |
| 10 | .20 .24 .08 .04 .00 .12 .20 .00 .32 .04 .04 .00 .00 .00 .36 |
| 11 | .80 .04 .04 .00 .08 .16 .00 .00 .00 .04 .00 .00 .00 .00 .00 |
| 12 | .00 .00 .00 .00 .04 .24 .00 .32 .08 .00 .04 .04 .16 .00 .00 |
| 13 | .68 .04 .32 .00 .08 .16 .08 .00 .68 .00 .08 .00 .00 .00 .00 |
| 14 | .52 .08 .04 .04 .08 .64 .28 .04 .24 .00 .12 .04 .00 .28 .16 |
| 15 | .00 .00 .00 .00 .08 .08 .12 .20 .12 .00 .00 .04 .12 .00 .00 |
| 16 | .24 .08 .08 .08 .00 .64 .24 .00 .48 .00 .12 .08 .00 .00 .28 |
| 17 | .52 .28 .04 .00 .04 .40 .32 .12 .20 .04 .12 .08 .04 .16 .04 |
| 18 | .00 .00 .00 .00 .08 .28 .04 .00 .12 .00 .20 .00 .04 .00 .00 |
| 19 | .00 .00 .00 .00 .08 .44 .12 .08 .12 .00 .16 .04 .00 .00 .00 |
| 20 | .00 .04 .00 .00 .04 .44 .16 .56 .00 .20 .00 .12 .40 .00 .00 |

| Face | Features 076 to 090 |
|---|---|
| 01 | .04 .04 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .04 .04 |
| 02 | .00 .08 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .04 .00 .04 |
| 03 | .00 .00 .44 .76 .76 .32 .40 .48 .52 .00 .04 .00 .08 .16 .24 |
| 04 | .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .24 .04 .00 |
| 05 | .16 .04 .00 .44 .00 .60 .40 .00 .00 .56 .20 .04 .00 .00 .04 |
| 06 | .00 .04 .00 .00 .00 .00 .00 .00 .00 .00 .04 .00 .00 .00 .04 |
| 07 | .28 .04 .00 .60 .00 .00 .00 .92 .00 .04 .16 .00 .12 .00 .00 |
| 08 | .20 .12 .44 .68 .04 .08 .00 .68 .08 .00 .00 .04 .20 .04 .52 |
| 09 | .08 .12 .00 .32 .00 .28 .12 .20 .04 .00 .00 .00 .00 .00 .24 |
| 10 | .12 .04 .00 .04 .00 .00 .00 .04 .00 .00 .00 .00 .48 .24 .84 |
| 11 | .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .04 |
| 12 | .24 .12 .24 .92 .16 .48 .32 .64 .00 .00 .12 .04 .00 .04 .12 |
| 13 | .04 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 |
| 14 | .32 .00 .00 .08 .00 .00 .00 .04 .00 .00 .00 .00 .32 .20 .72 |
| 15 | .12 .16 .44 .88 .08 .44 .12 .76 .00 .60 .44 .00 .00 .24 .04 |
| 16 | .08 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .28 .24 .68 |
| 17 | .28 .04 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .20 .44 .60 |
| 18 | .20 .00 .04 .92 .00 .16 .68 .72 .00 .44 .36 .00 .00 .12 .00 |
| 19 | .12 .12 .00 .52 .00 .20 .80 .00 .00 .00 .04 .00 .08 .04 .00 |
| 20 | .24 .00 .36 .72 .52 .00 .08 .92 .04 .00 .08 .00 .08 .16 .08 |

| Face | Features 091 to 105 |
|---|---|
| 01 | .12 .04 .00 .08 .00 .08 .04 .04 .04 .12 .00 .04 .04 .40 .20 |
| 02 | .08 .04 .04 .12 .00 .00 .04 .00 .04 .00 .00 .04 .00 .08 .08 |
| 03 | .28 .12 .00 .04 .08 .04 .72 .40 .56 .12 .76 .36 .76 .36 .76 |
| 04 | .08 .08 .08 .00 .00 .28 .00 .08 .00 .24 .04 .00 .04 .00 .04 |
| 05 | .00 .00 .00 .08 .04 .12 .20 .00 .12 .24 .04 .00 .08 .00 .04 |
| 06 | .20 .04 .00 .00 .04 .16 .20 .36 .04 .08 .04 .04 .64 .16 .20 |
| 07 | .00 .04 .00 .00 .00 .00 .04 .00 .00 .00 .04 .00 .08 .04 .04 |
| 08 | .28 .20 .40 .04 .32 .04 .04 .32 .00 .00 .12 .16 .44 .20 .00 |
| 09 | .28 .04 .08 .00 .04 .08 .44 .12 .48 .08 .04 .24 .60 .16 .28 |
| 10 | .36 .24 .28 .12 .52 .04 .20 .04 .00 .04 .08 .08 .48 .16 .04 |
| 11 | .00 .00 .00 .00 .00 .00 .04 .04 .00 .04 .04 .00 .16 .00 .00 |
| 12 | .24 .04 .00 .04 .04 .08 .36 .36 .40 .12 .44 .28 .60 .20 .44 |
| 13 | .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .00 .04 .08 .00 .04 |
| 14 | .28 .32 .44 .04 .24 .00 .16 .04 .00 .00 .00 .04 .28 .36 .00 |
| 15 | .32 .04 .00 .04 .00 .04 .16 .24 .32 .12 .56 .28 .44 .12 .04 |
| 16 | .40 .24 .40 .08 .32 .16 .52 .44 .32 .04 .12 .28 .72 .40 .20 |
| 17 | .56 .32 .12 .20 .04 .12 .20 .44 .12 .04 .32 .16 .48 .32 .28 |
| 18 | .12 .04 .00 .04 .04 .00 .00 .04 .00 .00 .00 .04 .00 .04 .00 |
| 19 | .12 .00 .00 .00 .00 .08 .00 .00 .00 .00 .00 .00 .04 .08 .04 |
| 20 | .16 .04 .00 .04 .04 .20 .32 .28 .24 .04 .64 .12 .72 .44 .76 |

| Face | Features 106 to 120 |
|---|---|
| 01 | .52 .16 .08 .00 .04 .00 .20 .00 .00 .32 .04 .64 .00 .00 .04 |
| 02 | .04 .00 .04 .00 .12 .08 .00 .00 .16 .28 .08 .04 .00 .08 .08 |
| 03 | .24 .20 .48 .20 .48 .04 .00 .04 .12 .04 .40 .08 .04 .08 .04 |
| 04 | .00 .00 .04 .00 .00 .00 .20 .00 .00 .00 .12 .04 .08 .04 .08 |
| 05 | .00 .00 .12 .04 .08 .08 .04 .00 .04 .20 .08 .00 .00 .00 .00 |
| 06 | .28 .20 .60 .00 .00 .08 .20 .00 .00 .00 .12 .00 .00 .04 .04 |
| 07 | .00 .00 .12 .00 .16 .08 .00 .00 .04 .00 .00 .00 .00 .04 .00 |
| 08 | .12 .16 .12 .00 .20 .12 .08 .00 .00 .24 .28 .00 .00 .12 .04 |
| 09 | .12 .28 .24 .08 .08 .12 .20 .00 .12 .04 .08 .00 .00 .04 .04 |
| 10 | .24 .20 .08 .00 .24 .08 .00 .00 .00 .28 .04 .00 .16 .08 .04 |
| 11 | .00 .08 .12 .00 .12 .04 .04 .00 .04 .16 .20 .00 .00 .04 .08 |
| 12 | .08 .04 .28 .08 .36 .04 .08 .08 .12 .16 .24 .00 .00 .04 .00 |
| 13 | .04 .00 .08 .00 .00 .04 .16 .00 .00 .16 .00 .00 .00 .04 .04 |
| 14 | .20 .12 .08 .00 .16 .00 .00 .08 .00 .60 .16 .56 .16 .04 .00 |
| 15 | .08 .00 .16 .00 .40 .12 .00 .00 .00 .04 .20 .84 .08 .04 .08 |
| 16 | .32 .24 .12 .16 .16 .12 .08 .00 .00 .60 .24 .08 .16 .08 .12 |
| 17 | .32 .04 .24 .16 .12 .12 .20 .00 .08 .24 .24 .64 .44 .04 .08 |
| 18 | .00 .00 .08 .04 .36 .24 .08 .00 .00 .12 .04 .88 .08 .00 .00 |
| 19 | .08 .08 .00 .00 .12 .12 .12 .00 .00 .32 .12 .12 .00 .00 .04 |
| 20 | .16 .24 .36 .48 .44 .08 .04 .08 .12 .16 .20 .60 .04 .12 .04 |

| Face | Features 121 to 125 |
|---|---|
| 01 | .00 .12 .24 .28 .04 |
| 02 | .04 .00 .20 .04 .00 |
| 03 | .04 .12 .16 .04 .04 |
| 04 | .00 .20 .04 .20 .00 |
| 05 | .48 .00 .12 .08 .00 |
| 06 | .04 .08 .08 .28 .04 |
| 07 | .52 .00 .16 .00 .00 |
| 08 | .04 .24 .32 .44 .00 |
| 09 | .12 .12 .08 .28 .16 |
| 10 | .00 .28 .32 .36 .00 |
| 11 | .04 .04 .20 .20 .12 |
| 12 | .20 .04 .24 .00 .04 |
| 13 | .16 .04 .28 .12 .08 |
| 14 | .00 .20 .16 .32 .00 |
| 15 | .32 .00 .16 .08 .08 |
| 16 | .08 .36 .20 .64 .00 |
| 17 | .00 .16 .12 .36 .12 |
| 18 | .56 .00 .12 .08 .04 |
| 19 | .60 .00 .16 .08 .04 |
| 20 | .04 .12 .24 .08 .12 |

# REFERENCES

Alvarez, G. & Fuentes, P. (1994). Recognition of facial expression in diverging socioeconomic levels. Brain and Cognition, 25, 235 - 239.

American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4'th Ed.). Washington, DC: Author.

Andreasen, N.C. & Carpenter, W.T.Jr. (1993). Diagnosis and classification of schizophrenia. Schizophrenia Bulletin, 19, 199 - 214.

Appelo, M.T., Woonings, F.M.J., van Nieuwenjuizen, C.J., Emmelkamp, P.M.G., Slooff, C.J., Louwerens, J.W. (1992). Specific skills and social competence in schizophrenia. Acta Psychiatrica Scandinavica, 85, 419 - 422.

Archer, J., Hay, D.C., Young, A.W. (1992). Face processing in psychiatric conditions. British Journal of Clinical Psychology, 31, 45 - 61.

Archer, J., Hay, D.C., Young, A.W. (1994). Movement, face processing, and schizophrenia: Evidence of a differential deficit in expression analysis. British Journal of Clinical Psychology, 33, 517 - 528.

Bachneff, S.A. (1991). Positron Emission Tomography and Magnetic Resonance Imaging: A review and a local circuit neurons hypo(dys)function hypothesis of schizophrenia. Biological Psychiatry, 30, 857 - 886.

Baddeley, A.D. (1983). Working memory. Philosophical Transactions of the Royal Society of London, B 302, 311 - 324.

Baddeley, A. (1992). Working memory. Science, 255, 556-559.

Balanova [computer program]. (no date). Adapted to personal computer by C.A. Creider, Department of Anthropology, University of Western Ontario. (Available from R.C. Gardner, Department of Psychology, University of Western Ontario, London, ON.)

Ball, S., Basatto, G.F., David, A.S., Jones, S.H., Hemsley, D.R., Pilowski, L.S., Costa, D.C., Ell, P.J., & Kerwin, R.W. (1998). Cognition functioning and $\text{GABA}_{(A)}$ / benztropine receptor binding in schizophrenia: $\text{A}^{1\ 2\ 3\text{I- Iomazenil}}$ SPECT Study [CD-ROM]. Biological Psychiatry, 43, 107 - 117. Abstract from WinSPIRS: PsychLit Item #1997 - 39188 - 004. Review available at http://www.futur.com/resframes.htm.

Bamber, D. (1979). State-trace analysis: A method of testing simple theories of causation. Journal of Mathematical Psychology, 19, 137 - 181.

Bamber, D. & van Santen, J.P.H. (1985). How many parameters can a model have and still be testable? Journal of Mathematical Psychology, 29, 443 - 473.

Bamber, D. & van Santen, J.P.H. (2000). How to assess a model's testability and identifiability. Journal of Mathematical Psychology, 44, 20 - 40.

Barta, P.E., Pearlson, G.D., Powers, R.E., Richards, S.S., & Tune, L.E. (1990). Auditory hallucinations and smaller superior temporal gyral volume in schizophrenia. American Journal of Psychiatry, 147, 1457 - 1462.

Bartha, R.B., Williamson, P.C., Drost, D.J., Malla, A., Carr, T.J., Cortese, L., Canaran, G., Rylett, R.J., & Neufeld, R.W.J. (1997). Measurement of glutamate and gluatmine in the medial prefrontal cortex of never-treated schizophrenic patients and healthy controls by proton Magnetic Resonance Spectroscopy. Archives of General Psychiatry, 54, 959 - 965.

Beck, A.T. (1967). Depression: Clinical, experimental, and theoretical aspects. New York, NY: Harper and Row.

Beech, A., McManus, D., Baylis, G., Tipper, S., Agar, K. (1991). Individual differences in cognitive processes: Towards an explanation of schizophrenic symptomatology. British Journal of Psychology, 82, 417 - 426.

Bellack, A.S., Blanchard, J.J., & Mueser, K.T. (1996). Cue availability and affect perception in schizophrenia. Schizophrenia Bulletin, 22, 535 - 544.

Bellack, A.S., Mueser, K.T., Wade, J., Sayers, S., & Morrison, R.L. (1992). The ability of schizophrenics to perceive and cope with negative affect. British Journal of Psychiatry, 160, 473 - 480.

Benes, F.M., Davidson, J., & Bird, E.D. (1986). Quantitative cytoarchitectural studies of the cerebral cortex of schizophrenics. Archives of General Psychiatry, 43, 31 - 35.

Benes, F.M., McSparren, J., Bird, E.D., SanGiovanni, J.P., & Vincent, S.L. (1991). Deficits in small interneurons in prefrontal and cingulate cortices of schizophrenic and schizoaffective patients. Archive of General Psychiatry, 48, 996 - 1001.

Benson, D.F. & Stuss, D.T. (1990). Frontal lobe influences on delusions: A clinical perspective. Schizophrenia Bulletin, 16, 403 - 411.

Bloom, F.E., Schulman, J.A., & Koob, G.F. (1989). Catecholamines and behaviour. In U. Trendelenburg & N. Weiner (Eds.), Heffter-Heubner's Handbook of Experimental Pharmacology 90: Catecholamines II (pp. 27 - 88). Heidelberg, Germany: Springer-Verlag.

Borod, J.C., Alpert, M., Brozgold, A., Martin, C., Welkowitz, J., Diller, L., Peselow, E., Angrist, B., & Lieberman, A. (1989). A preliminary comparison of flat affect schizophrenics and brain-damaged patients on measures of affective processing. Journal of Communication Disorders, 22, 93 - 104.

Borod, J.C., Koff, E., Lorch, M.P., & Nicholas, M. (1986). The expression and perception of facial emotion in brain-damaged patients. Neuropsychologia, 24, 169 - 180.

Browne, M.W. (2000). Cross-validation methods. Journal of Mathematical Psychology, 44, 108 - 132.

Bryson, G., Bell, M., & Lysaker, P. (1997). Affect recognition in schizophrenia: A function of global impairment or a specific cognitive deficit. Psychiatry Research, 71, 105 - 113.

Buffery, A.W.H. & Burton, A. (1982). Information processing and redevelopment: Towards a science of neuropsychological rehabilitation. In A. Burton (Ed.), The pathology and psychology of cognition (pp. 253 - 292). London, UK: Methuen.

Burch, J.W. (1995). Typicality range deficit in schizophrenic's recognition of emotion in faces. Journal of Clinical Psychology, 51, 140 - 152.

Busemeyer, J.R. & Townsend, J.T. (1993). Decision field theory: A dynamic cognitive approach to decision making in an uncertain environment. Psychological Review, 100, 432 - 459.

Canadian Pharmacists Association (1998). Compendium of pharmaceuticals and specialties (33'rd Ed.). Ottawa, ON: Author.

Carey, S. (1992). Becoming a face expert. Philosophical Transactions of the Royal Society of London, Series B, 335, 95 - 103.

Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition – A maturational component? Developmental Psychology, 16, 257 - 269.

Carlsson, M. & Carlsson, A. (1990). Schizophrenia: A subcortical neurotransmitter imbalance syndrome? Schizophrenia Bulletin, 16, 425 - 432.

Carpenter, W.T.Jr., Buchanan, R.W., Kirkpatrick, B., Tamminga, C., & Wood, F. (1993). Strong inference, theory testing, and the neuroanatomy of schizophrenia. Archives of General Psychiatry, 50, 825 - 831.

Carroll, J.M. & Russell, J.A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. Journal of Personality and Social Psychology, 70, 205 - 218.

Carroll, J.M. & Russell, J.A. (1997). Facial expressions in Hollywood's portrayal of emotion. Journal of Personality and Social Psychology, 72, 164 - 176.

Carter, J.R. (1994). Encoding for similarity judgements among paranoid and nonparanoid schizophrenics. Unpublished Master's thesis. University of Western Ontario: London, ON.

Carter, J.R. (1997 - 1999). JeffSim 1.0 [Computer program]. London, ON: University of Western Ontario, Department of Psychology.

Carter, J.R. (2000). Facial expression analysis in schizophrenia (Unabridged version). Unpublished manuscript. (Available from the author.)

Carter, J.R. & Neufeld, R.W.J. (1996a). A process model of salience deficits in multidimensional judgments (Research Bulletin 738). London, Ontario: University of Western Ontario, Department of Psychology.

Carter, J.R. & Neufeld, R.W.J. (1996b). Similarity judgments of words and facial expressions in schizophrenia. Poster session presented at the convention of the American Psychological Association, Toronto, Ontario, Canada.

Carter, J.R. & Neufeld, R.W.J. (1999). Cognitive processing of multidimensional stimuli in schizophrenia: Formal modeling of judgment speed and content. Journal of Abnormal Psychology, 108, 633 - 654.

Carter, J.R., Neufeld, R.W.J., & Benn, K. (1998). Application of process models in assessment psychology: Potential assets and challenges. Psychological Assessment, 10, 379 - 395.

Cegalis, J.A., Leen, D., & Solomon, E.J. (1977). Attention in schizophrenia: An analysis of selectivity in the functional visual field. Journal of Abnormal Psychology, 86, 470 - 482.

Chandler, J.P. (1975). STEPIT 7.4 [computer program]. Stillwater, OK: Oklahoma State University, Computer Science Department.

Chapman, L.J. & Chapman, J.P. (1973). Problems in the measurement of disordered thought. Disordered thought in schizophrenia (pp. 53 - 76). New York, NY: Appleton-Century-Crofts.

Chechile, R.A. (1987). Trace susceptibility theory. Journal of Experimental Psychology: General, 116, 203 - 222.

Cicchetti, D. & Cohen, D.J. (1995). Developmental psychopathology. New York: NY: Wiley.

Clark, L.A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. Annual Review of Psychology, 46, 121 - 153.

Crick, F. (1989). The recent excitement about neural networks. Nature, 337, 129 - 132.

Cohen, J.D. & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. Psychological Review, 99, 45 - 77.

Cole, J. (1998). About face. Cambridge, MA: MIT Press.

Cornblatt, B.A., Lenzenweger, M.F., Dworkin, R.H., & Erlenmeyer-Kimling, L. (1992). Childhood attentional dysfunctions predict social deficits in unaffected adults at risk for schizophrenia. British Journal of Psychiatry, 161 (suppl. 18), 59 - 64.

Cornblatt, B.A., Lanzenwerger, M.F., Erlenmeyer-Kimling, L. (1989). The Continuous Performance Test, Identical Pairs version (CPT-IP): II. Contrasting attentional profiles in schizophrenic and depressed patients. Psychiatry Research, 29, 65 - 85.

Cornblatt, B.A., Risch, N.J., Faris, G., Friedman, D., & Erlenmeyer-Kimling, L. (1988). The Continuous Performance Test, Identical Pairs version (CPT-IP): I. New findings about sustained attention in normal families. Psychiatry Research, 26, 223 - 238.

Corrigan, P.W. & Green, M.F. (1993). Schizophrenic patients' sensitivity to social cues: The role of abstraction. American Journal of Psychiatry, 150, 589 - 594.

Corrigan, P.W., Silverman, R., Stephenson, J., Nugent-Hirschbeck, J., & Buican, B.J. (1996). Situational familiarity and feature recognition in schizophrenia. Schizophrenia Bulletin, 22, 153 - 161.

Cottrell, G.W., Dailey, M.N., Padgett, C., & Adophs, R. (in press). Is all face processing holistic? The view from UCSD. In M. Wenger, & J. Townsend (Eds.), Computational, geometric, and process perspectives on facial cognition: Contexts and challenges. Erlbaum.

Cramer, P., Weegmann, M., O'Neil, M. (1989). Schizophrenia and the perception of emotions: How accurately do schizophrenics judge the emotional states of others. British Journal of Psychiatry, 155, 225 - 228.

Creese, I., Burt, D.R., & Snyder, S.H. (1976). Dopamine receptor binding predicts clinical and pharmacological potencies of antipsychotic drugs. Science, 192, 481 - 483.

Crow, T.J. (1990). Temporal lobe asymmetries as the key to the etiology of schizophrenia. Schizophrenia Bulletin, 16, 433 - 443.

Cutting, J. (1981). Judgement of emotional expression in schizophrenia. British Journal of Psychiatry, 139, 1 - 6.

David, A.S. (1994). Dysmodularity: A neurological model for schizophrenia. Schizophrenia Bulletin, 20, 249 - 255.

Davis, J.A. (1970). Some pitfalls of data analysis without a formal criterion. In D.E. Morrison & R.E. Henkel (Eds.), The significance test controversy: A reader (pp. 91 - 93). Chicago, IL: Aldine.

Davis, J.M. (1976). Comparative doses and costs of antipsychotic medication. Archives of General Psychiatry, 33, 858 - 861.

Deakin, J.F.W. (1994). Neuropsychological implications of brain changes in schizophrenia: An overview. Psychopathology, 27, 251 - 254.

Deakin, J.F.W., Slater, P., Simplson, M.D.C., Gilchrist, A.C., Skan, W.J., Royston, M.C., Reynolds, G.P., & Cross, A.J. (1989). Frontal cortical and left temporal glutamatergic dysfunction in schizophrenia. Journal of Neurochemistry, 52, 1781 - 1786.

Deering, G. (1963). Affective stimuli and disturbance of thought processes. Journal of Consulting Psychology, 27, 338 - 343.

Deutsch, S.I., Mastropaolo, J. Schwartz, B.L., Rosse, R.B., & Morihisa, J.M. (1989). A "glutamatergic hypothesis" of schizophrenia. Clinical Neuropharmacology, 12, 1 - 13.

Devlin, J.T., Gonnerman, L.M., Andersen, E.S., & Seidenberg, M.S. (1998). Category-specific semantic deficits in focal and widespread brain damage: A computational account. Journal of Cognitive Neuroscience, 10, 77 - 94.

Dimberg, U. & Öhman, A.(1996). Behold the wrath: Psychophysiological responses to facial stimuli. Motivation and Emotion, 20, 149 - 182.

Dougherty, F.E., Bartlett, E.S., & Izard, C.E. (1974). Responses of schizophrenics to expressions of the fundamental emotions. Journal of Clinical Psychology, 30, 243 - 246.

Dunn, L.M. & Dunn, L.M. (1981). Peabody Picture Vocabulary Test – Revised. Circle Pines, MN: American Guidance Service.

Edelman, G.M. (1987). Neural Darwinism: The theory of neuronal group selection. New York, NY: Basic Books.

Edwards, K. (1998). The face of time: Temporal cues in facial expressions of emotion. Psychological Science, 9, 270 - 276.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. Psychological Bulletin, 115 (2), 268 - 287.

Ekman, P. & Friesen, W.V. (1975). Unmasking the face: A guide to recognizing emotions from facial expressions. Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P. & Friesen, W.V. (1976). Pictures of facial affect. Palo Alto, CA: Consulting Psychologists Press.

Ellis, H.D. (1992). The development of face processing skills. In V. Bruce & A. Cowey (Eds.), Processing the facial image (pp. 105 - 111). Oxford, UK: Oxford University Press.

Elman, J.L. (1990). Finding structure in time. Cognitive Science, 14, 179 - 211.

Ennis, J.H., & Whelton, C. (1994). The relationship between face recognition, facial affect recognition and social skills in schizophrenia. Canadian Journal of Psychiatry, 39, 58.

Erwin, R.J., Gur, R.C., Gur, R.E., Skolnick, B., Mawhinney-Hee, M., & Smailis, J. (1992). Facial emotion discrimination: I. Task construction and behavioral findings in normal subjects. Psychiatry Research, 42, 231 - 240.

Farah, MJ. & McClelland, J.L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specifity. Journal of Experimental Psychology: General, 120, 339 - 357.

Farde, L., Pauli, S., Hall, H., Eriksson, L., Halldin, C., Högberg, T., Nilsson, L, Sjögren, I., & Stone-Elander, S. (1988). Stereoselective binding of 11C-raclopride in living human brain -- a search for extrastriatial central D2-dopamine receptors by PET. Psychopharmacology, 94, 471 - 478.

Feinberg, I. (1982 / 1983). Schizophrenia: Caused by a fault in programmed synaptic elimination during adolescence? Journal of Psychiatric Research, 17, 319 - 334.

Feinberg, I., (1990). Cortical pruning and the development of schizophrenia. Schizophrenia Bulletin, 16, 567 - 568.

Feinberg, T.E., Rifkin, A., Schaffer, C., & Walker, E. (1986). Facial discrimination and emotional recognition in schizophrenia and affective disorders. Archives of General Psychiatry, 43, 276 - 279.

Flack, W.F.Jr., Cavallaro, L.A., Laird, J.D., & Miller, D.R. (1997). Accurate encoding and decoding of emotional facial expressions in schizophrenia. Psychiatry, 60, 197 - 210.

Fonnum, F. (1984). Glutamate: A neurotransmitter in mammalian brain. Journal of Neurochemistry, 42, 1 - 11.

Foote, S.L., Freedman, R., & Oliver, A.P. (1975). Effects of putative neurotransmitters on neuronal activity in monkey auditory cortex. Brain Research, 86, 229 - 242.

Fox, E., Lester, V., Russo, R., Bowles, R.J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? Cognition and Emotion, 14, 61 - 92.

Fridlund, A.J. (1992). Darwin's anti-Darwinism in the Expression of the emotions in man and animals. In K.T. Strongman (Ed.), International review of studies on emotion, Vol. 2 (pp. 117 - 137). New York, NY: Wiley.

Fried, I., Mateer, C., Ojemann, G., Wohns R., & Fedio, P. (1982). Organization of visuospatial functions in human cortex: Evidence from electrical stimulation. Brain, 105, 349 - 371.

Friedman, D., Vaughan, H.G., Erlenmeyer-Kimling, L. (1981). Multiple late positive potentials in two visual discrimination tasks. Psychophysiology, 18, 635 - 649.

Frith, C.D. (1979). Consciousness, information processing and schizophrenia. British Journal of Psychiatry, 134, 225 - 235.

Frith, C.D., Stevens, M., Johnstone, E.C., Owens, D.G.C., & Crow, T.J. (1983). Integration of schematic faces and other complex objects in schizophrenia. Journal of Nervous and Mental Disease, 171, 34 - 39.

Gaebel, W. & Wölwer, W. (1992). Facial expression and emotional face recognition in schizophrenia and depression. European Archives of Psychiatry and Clinical Neuroscience, 242, 46 - 52.

Garety, P.A., Fowler, D., & Kuipers, E. (2000). Cognitive-behavioral therapy for medication-resistant symptoms. Schizophrenia Bulletin, 26, 73 - 86.

George, M.S., Ketter, T.A., Gill, D.S., Haxby, J.V., Ungerleider, L.G., Herscovitch, P., & Post, R.M. (1993). Brain regions involved in recognizing facial emotion or identiy: An oxygen-15 PET study. Journal of Neuropsychiatry and Clinical Neurosciences, 5, 394 - 394.

Gessler, S., Cutting, J., Frith, C.D., & Weinman, J. (1989). Schizophrenic inability to judge facial emotion: A controlled study. British Journal of Clinical Psychology, 28, 19 - 29.

Goldman-Rakic, P.S. (1987). Development of cortical circuitry and cognitive function. Child Development, 58 (3), 601-622.

Grant, D.A. (1962). Testing the null hypothesis and the strategy and tactics of investigating theoretical models. Psychological Review, 69, 54 - 61.

Green, M.F., Kern, R.s., Braff, D.L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the right stuff? Schizophrenia Bulletin, 26, 119 - 136.

Grossberg, S. (in press-a). How hallucinations may arise from brain mechanisms of learning, attention, and volition. Journal of the International Neuropsychological Society.

Grossberg, S. (in press-b). The imbalanced brain: From normal behavior to schizophrenia. Biological Psychiatry.

Harizuka, S. (1977). [Perception of schizophrenic patients in classifying pictures of facial expressions]. Japanese Journal of Psychology, 48, 231 -238.

Harris, J.C. (1995). Schizophrenia: A neurodevelopmental disorder. Developmental Neuropsychiatry: Assessment, diagnosis, and treatment of developmental disorders (Vol. 2, pp. 404 - 426). New York, NY: Oxford University Press.

Hasselmo, M.E., Rolls, E.T., & Baylis, G.C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. Behavioural Brain Research, 32, 203 - 218.

Hebb, D.O. (1949). The organization of behavior: A neuropsychological theory. New York, NY: Wiley.

Heimberg, C., Gur, R.E., Erwin, R.J., Shtasel, D.L., & Gur, R.C. (1992). Facial emotion discrimination: III. Behavioral findings in schizophrenia. Psychiatry Research, 42, 253 - 265.

Heit, G., Smith, M.E., & Halgren, E. (1988). Neural encoding of individual words and faces by the human hippocampus and amygdala. Nature, 333, 773 - 775.

Hemsley, D.R. (1982). Cognitive impairment in schizophrenia. In A. Burton (Ed.), The pathology and psychology of cognition (pp. 169 - 203). London, UK: Methuen.

Heywood, C.A. & Cowey, A. (1992). The role of the "face-cell" area in the discrimination and recognition of faces by monkeys. Philosophical Transactions of the Royal Society of London, Series B, 335, 31 - 38.

Hinton, G.E. & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. Psychological Review, 98, 74 - 95.

Hoffman, R.E. (1992). Attractor networks and psychotic disorders [CD-ROM]. Psychiatric Annals, 27, 119 - 124. Abstract from PsychInfo Item: 1992-33482-001.

Hoffman, R.E. & Dobscha, S.K. (1989). Cortical pruning and the development of schizophrenia. Schizophrenia Bulletin, 15, 477 - 490.

Hoffman, R.E. & Dobscha, S.K. (1990). Reduced cortico-cortical exchange of information as a possible etiology of schizophrenia: Reply to Feinberg. Schizophrenia Bulletin, 16, 569 - 570.

Hoffman, R.E. & McGlashan, T.H. (1993). Parallel distributed processing and the emrgence of schizophrenic symptoms. Schizophrenia Bulletin, 19, 119 - 140.

Hoffman, R.E. & McGlashan, T.H. (1994). Corticocortico connectivity, autonomous networks, and schizophrenia. Schizophrenia Bulletin, 20, 257 - 261.

Hoffman, R.E. & McGlashan, T.H. (1998). Reduced corticocortico connectivity can induce speech perception pathology and hallucinated ‘voices.’ Schizophrenia Research, 30, 137 - 141.

Hoffman, R.E., Rapaport, J., Ameli, R., McGlashan, T.H., Harcherik, D., & Servan-Schreiber, D. (1995). A neural network simulation of hallucinated “voices” and associated speech perception impairments in schizophrenic patients. Journal of Cognitive Neuroscience, 7, 479 - 496.

Hollingshead, B.A. (1957). Two factor index of social position. Unpublished manuscript. (Available from the current author.)

Hooker, C., Park, S., & Tajuddin, Z. (1999). Eye gaze direction perception in schizophrenia patients. Paper presented to the Fourteenth Annual Meeting of the Society for Research in Psychopathology, Montreal, Quebec, Canada.

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

Hughes, H.C., Reuter-Lorenz, P.A., Nozawa, G., & Fendrich, R. (1994). Visual-auditory interactions in sensorimotor processing: Saccades versus manual responses. Journal of Experimental Psychology: Human Perception and Performance, 20, 131 - 153.

Hunter, J.E., Harris, R.J., & Wilkinson, L. (1996). Significance tests: Should they be banned from APA Journals? Symposium presented at the 104'th Annual Convention of the American Psychological Association. Toronto: ON. [Available on cassette.]

Hush, D.R., Horne, B., & Salas, J.M. (1992). Error surfaces for multilayer perceptrons. IEEE Transactions on Systems, Man, and Cybernetics, 22, 1152 - 1161.

Hüttenlocher, P.R. (1979). Synaptic density in human frontal cortex: Developmental changes and effects of aging. Brain Research, 163, 195 - 205.

Hüttenlocher, P.R. (1984). Synapse elimination and plasticity in developing human cerebral cortex. American Journal of Mental Deficiency, 88 (5), 488 - 496.

Hüttenlocher, P.R., deCourten, C., Garey, L.J., & van der Loos, H. (1982) synaptogenesis in human visual cortex: Evidence for synapse elimination during normal development. Neuroscience Letters, 33, 247 - 252.

Ingvar, D.H. & Franzén, G. (1974). Abnormalities of cerebral blood flow distribution in patients with chronic schizophrenia. Acta Psychiatrica Scandinavica, 50, 425 - 462.

Iscoe, I. & Veldman, D.J. (1963). Perception of an emotional continuum by schizophrenics, normal adults, and children. Journal of Clinical Psychology, 19, 272 - 276.

Izard, C.E. (1959). Paranoid schizophrenic and normal subjects' perceptions of photographs of human faces. Journal of Consulting Psychology, 23, 119 - 124.

Izard, C.E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. Psychological Bulletin, 115, 288 - 299.

Jeste, D.V., McAdams, La., Palmer, B.W., Braff, D., Jernigan, T.L., Paulsen, J.S., Stout, J.C., Symonds, L.L., Bailey, A., & Heaton, R.K. (1998). Relationship of neuropsychological and MRI measures to age of onset of schizophrenia. Acta Psychiatrica Scandinavica, 98, 156 - 164.

Jetté, J.L. (in preparation). A six-dimensional dynamical system of psychological stress and coping. Unpublished Doctoral dissertation. London, ON: University of Western Ontario.

Johnstone, E.C., MacMillan, J.F., Frith, C.D., Benn, D.K., & Crow, T.J. (1990). Further investigation of the predictors of outcome following first schizophrenic episodes. British Journal of Psychiatry, 157, 182 - 189.

Jordan, M.I. (1992). Serial order: A parallel distributed processing approach (Technical Report 8604). San Diego, CA: Institute for Cognitive Science, University of California.

Jordan, M.I. & Rumelhart, D.E. (1992). Forward models: Supervised learning with a distal teacher. Cognitive science, 16, 307 - 354.

Kalat, J.W. (1988). Biological Psychology (3'rd Ed.). Belmont, CA: Wadsworth.

Kapur, S., Zipursky, R.B., Remington, G., Jones, C., DaSilva, J., Wilson, A.A., & Houle, S. (1998). 5-$HT_2$ Receptor occupancy of olanzapine in schizophrenia: A PET investigation. American Journal of Psychiatry, 155, 921 - 928.

Katsikitis, M. (1997). The classification of facial expressions of emotion: A multidimensional-scaling approach. Perception, 26, 613 - 626.

Kee, K.S., Kern, R.S., Marshall, B.D.Jr., & Green, M.F. (1998). Risperidone versus haloperidol for perception of emotion in treatment-resistant schizophrenia: Preliminary findings. Schizophrenia Research, 31, 159 - 165.

Kendall, R.W. & Gourlay, J. (1970). The clinical distinction between the affective psychoses and schizophrenia. British Journal of Psychiatry, 117, 261 - 266.

Kern, R.S., Green, M.F., Marshall, B.D.Jr., Wirshing, W.C., Wirshing, D., McGurk, S., Marder, S.R., & Mintz, J. (1998). Risperidone vs. Halperidol on reaction time, manual dexterity, and motor learning in treatment-resistant schizophrenia patients. Biological Psychiatry, 44, 726 - 732.

Kerr, S.L. & Neale, J.M. (1993). Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? Journal of Abnormal Psychology, 102, 312 - 318.

Kimura, D. (1986). Neuropsychology test procedures. London, ON: DK Consultants.

Kline, J.S., Smith, J.E., & Ellis, H.C. (1992). Paranoid and nonparanoid schizophrenic processing of facially displayed affect. Journal of Psychiatric Research, 26, 169 - 182.

Knight, R.A. & Silverstein, S.M. (1998). The role of cognitive psychology in guiding research on cognitive deficits in schizophrenia: A process-oriented approach. In M.F. Lenzenweger & R.H. Dworkin (Eds.), Origins and development of schizophrenia: Advances in experimental psychopathology (pp. 247 - 295). Washington, DC: American Psychological Association.

Kokkinidis, L. & Anisman, H. (1980). Amphetamine models of paranoid schizophrenia: An overview and elaboration of animal experimentation. Psychological Bulletin, 88, 551 - 579.

Kotrla, K.J. & Weinberger, D.R. (1995). Brain imaging in schizophrenia. Annual Review of Medicine, 46, 113 - 122.

LaRusso, L. (1978). Sensitivity of paranoid patients to nonverbal cues. Journal of Abnormal Psychology, 87, 463 - 471.

Lay, K.L., Waters, E., & Park, K.A. (1989). Maternal responsiveness and child compliance: The role of mood as a mediator. Child Development, 60, 1405 - 1411.

Ledingham, J.E. (1981). Developmental patterns of aggressive and withdrawn behavior in childhood: A possible method for identifying preschizophrenics. Journal of Abnormal Child Psychology, 9, 1 - 22.

Ledingham, J.E. & Schwartzman, A.E. (1984). A 3-year follow-up of aggressive and withdrawn behavior in childhood: Preliminary findings. Journal of Abnormal Child Psychology, 12, 157 - 168.

Lee, T. & Seeman, P. (1980). Elevation of brain neuroleptic / dopomine receptors in schizophrenia. American Journal of Psychiatry, 137, 191 - 197.

Levy, L.H., Orr, T.B., Rosenzweig, S. (1960). Judgments of emotion from facial expressions by college students, mental retardates, and mental hospital patients. Journal of Personality, 28, 342 - 349.

Lewis, S.F. & Garver, D.L. (1995). Treatment and diagnostic subtype in facial affect recognition in schizophrenia. Journal of Psychiatric Research, 29, 5 - 11.

Li, S.-C. (1994). A diagnostic approach to model evaluation: Parameter sensitivity and interdependence analysis in models of cognition. Unpublished doctoral dissertation. Norman, OK: University of Oklahoma.

Long, D.L., Parks, R.W., & Levine, D.S. (1998). An introduction to neural network modeling: Merits, limitations, and controversies. In R.W. Parks, D.S. Levine, and D.L. Long (Eds.), Fundamentals of neural network modeling: Neuropsychology and cognitive neuroscience (pp. 3 - 31). Cambridge, MA: MIT Press.

Lubow, R.E., Weiner, I., Schlossberg, A., & Baruch, I. (1987). Latent inhibition and schizophrenia. Bulletin of the Psychonomic Society, 25, 464 - 476.

Luce, R.D. (1986). Response times: Their role in inferring elementary mental organization. New York, NY: Oxford University Press.

Magaro, P.A., Abrams, L., & Cantrell, P. (1981). The Maine scale of paranoid and nonparanoid schizophrenia: Reliability and validity. Journal of Consulting and Clinical Psychology, 49, 438 - 447.

Maher, B.A., Manschreck, T.C., Woods, B.T., Yurgelun-Todd, D.A., & Tsuang, M.T. (1995). Frontal lobe volume and context effets in short-term recall in schizophrenia. Biological Psychiatry, 37, 144 - 150.

Mandal, M.K. & Gewali, H. (1989). Identification of brief presentations of facial expressions of affects in schizophrenics. International Journal of Psychology, 24, 605 - 616.

Mandal, M.K. & Rai, A. (1987). Responses to facial emotion and psychopathology. Psychiatry Research, 20, 317 - 323.

Manstead, A.S.R. & Edwards, R. (1992). Communicative aspects of children's emotional competence. In K.T. Strongman (Ed.), International review of studies on emotion, Vol. 2 (pp. 167 - 95). New York, NY: Wiley.

Marr, D. (1982). Vision. San Francisco, CA: Freeman.

MathWorks, Inc. (1984 - 1999) MATLAB: The language of technical computing [computer program]. Natick, MA: Author.

Mazzoni, P., Andersen, R.A., Jordan, M.I. (1991). A more biologically plausible learning rule than backpropagation applied to a network model of cortical area 7a [CD-ROM]. Cerebral Cortex, 1, 293 - 307. Abstract from: SilverPlatter: PsychLIT Item: 80-20564.

McAlpine, C., Singh, N.N., Ellis, C.R., Kendall, K.A., & Hampton, C. (1992). Enhancing the ability of adults with mental retardation to recognize facial expressions of emotion. Behavior Modification, 16, 559 - 573.

McCloskey, M. (1991). Networks and theories: The place of connectionism in cognitive science. Psychological Science, 2, 387 - 395.

McCracken, G.D. (1988). The long interview. Newbury Park, CA: Sage.

McFall, R.M., Treat, T.A., & Viken, R.J. (1997). Contributions of cognitive theory to new behavioral treatments. Psychological Science, 8, 174 - 176.

McRae, K., & Hetherington, P.A. (1993). Catastrophic interference is eliminated in pretrained networks. In Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society (pp. 723 - 728). Hillsdale, NJ: Erlbaum.

Medalia, A., Aluma, M., Tyron, W., & Merriam, A.E. (1998). Effectiveness of attention training in schizophrenia. Schizophrenia Bulletin, 24, 147 - 152.

Meehl, P.E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. Journal of Consulting and Clinical Psychology, 46, 806 - 834.

Meehl, P.E. (1983). Consistency tests in estimating the completeness of the fossil record: A neo-Popperian approach to statistical paleontology. In J. Earman (Ed.), Minnesota studies in the philosophy of science: Volume X: Testing scientific theories (pp. 413 - 473). Minneapolis, MN: University of Minnesota Press.

Meehl, P.E. & Golden, R.R. (1982). Taxometric methods. In P.C. Kendall & J.N. Butcher (Eds.), Handbook of research methods in clinical psychology (pp. 127 - 181). New York, NY: Wiley.

Microsoft Corporation (1994 - 1997). Microsoft Developer Studio 97, Professional Edition: Visual C++ 5.0 [Computer program]. Redman, WA: Microsoft Corporation.

Milner, B. (1963). Effects of different brain lesions on card sorting. Archives of Neurology, 9, 90 - 100.

Minsky, M. & Papert, S. (1969). Perceptrons: An introduction to computational geometry. Cambridge, MA: MIT Press.

Morrison, D.E. & Henkel, R.E. (Eds.) (1970). The significance test controversy. Chicago, IL: Aldine.

Morrison, R.L. (1991). Schizophrenia. In M. Hersen & S.M. Turner (Eds.), Adult psychopathology and diagnosis (2'nd Ed., pp. 149 - 169). New York, NY: Wiley.

Morrison, R.L., Bellack, A.S., Bashore, T.R. (1988). Perception of emotion among schizophrenic patients. Journal of Psychopathology and Behavioral Assessment, 10, 319 - 332.

Morrison, R.L, Bellack, A.S., & Mueser, K.T. (1988). Deficits in facial-affect recognition and schizophrenia. Schizophrenia Bulletin, 14, 67 - 83.

Mueser, K.T., Bellack, A.S., Morrison, R.L, & Wade, J.H. (1990). Gender, social competence, and symptomatology in schizophrenia: A longitudinal analysis. Journal of Abnormal Psychology, 99, 138 - 147.

Mueser, K.T., Doonan, R., Penn, D.L., Blanchard, J.J., Bellack, A.S., Nishith, P., & DeLeon, J. (1996). Emotion recognition and social competence in chronic schizophrenia. Journal of Abnormal Psychology, 105, 271 - 275.

Mueser, K.T., Penn, D.L., Blanchard, J.J., & Bellack, A.S. (1997). Affect recognition in schizophrenia: A synthesis of findings across three studies. Psychiatry, 60, 301 - 308.

Muliak, S.A., James, L.R., VanAlstine, J., Bennett, N., Lind, S., & Stillwell, C.D., (1989). Evaluation of goodness-of-fit indices for structural equation models. Psychological Bulletin, 105, 430 - 445.

Muzekari, J.H. & Bates, M.E. (1977). Judgment of emotion among chronic schizophrenics. Journal of Clinical Psychology, 33, 662 - 666.

Myers, J.L, & Well, A.D. (1991). Research design and statistical analysis. New York, NY: HarperCollins.

Nelson, P.G., Fields, R.D., Yu, C., & Neale, E.A. (1990). Mechanisms involved in activity-dependent synapse formation in mammalian central nervous system cell cultures. Journal of Neurobiology, 21, 138 - 156.

Neufeld, R.W.J. (1976). Simultaneous processing of multiple stimulus dimensions among paranoid and nonparanoid schizophrenics. Multivariate Behavioral Research, 4, 425 - 442.

Neufeld, R.W.J. (1991). Memory in paranoid schizophrenia. In P.A. Magaro (Ed.), Cognitive bases of mental disorders (pp. 231 - 261). Newbury Park, CA: Sage.

Neufeld, R.W.J., Carter, J.R., Nicholson, I.R., & Vollick, D. (1999). Schizophrenia. In W.L. Marshall & P. Firestone (Eds.), Abnormal psychology perspectives (pp. 344 - 372). Scarborough, ON: Prentice Hall.

Neufeld, R.W.J., Vollick, D., & Highgate, S. (1993). Stochastic modeling of stimulus encoding and memory search in paranoid schizophrenia: Clinical and theoretical implications. In R.L. Cromwell & C.R. Snyder (Eds.), Schizophrenia: Origins, processes, treatment, and outcome (pp. 176 - 196). New York, NY: Oxford.

Neufeld, R.W.J., & Williamson, P. (1996). Neuropsychological correlates of positive symptoms: Delusions and hallucinations. In C. Pantelis, H.E. Nelson, & T.R.E. Barnes (Eds.), Schizophrenia: A neuropsychological perspective (pp. 205-235). London: Wiley.

Nicholson, I.R. & Neufeld, R.W.J. (1993). The classification of the schizophrenias according to symptomatology: A two-factor model. Journal of Abnormal Psychology, 102, 259 - 270.

Nolte, J. (1993). The human brian: An introduction to its functional anatomy (3'rd Ed.). St. Louis, MO: Mosby-Year Book.

Novic, J., Luchins, D.J., & Perline, R. (1984). Facial affect recognition in schizophrenia: Is there a differential deficit? British Journal of Psychiatry, 144, 533 - 537.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia, 9, 97 - 113.

Ortony, A. & Turner, T.J., (1990). What's so basic about basic emotions? Psychological Review, 97, 315 - 331.

Pachella, R.G. (1974). The interpretation of reaction time in information-processing research. In B.H. Kantowitz (Ed.), Human information processing: Tutorials in performance and cognition (pp. 41 - 82). Hillsdale, NJ: Erlbaum.

Parks, R.W. & Levine, D.S. (1998). Neural network modeling of Wisconsin Card Sorting and Verbal Fluency tests: Applications with frontal lobe-damaged and Alzheimer's disease patients. In R.W. Parks, D.S. Levine, and D.L. Long (Eds.), Fundamentals of neural network modeling: Neuropsychology and cognitive neuroscience (pp. 357 - 380). Cambridge, MA: MIT Press.

Penn, D.L., Corrigan, P.W., Bentall, R.P., Racenstein, J.M., & Newman, L. (1997). Social cognition in schizophrenia. Psychological Bulletin, 121, 114 - 132.

Penn, D.L. & Mueser, K.T. (1996). Research update on the psychosocial treatment of schizophrenia. American Journal of Psychiatry, 153, 607 - 617.

Penn, D.L. & Spaulding, P.W. (1997). Introduction: Factors underlying social functioning in schizophrenia: Information processing and social perception. Psychiatry, 60, 279 - 280.

Penn, D.L., van der Does, A.J.W., Spaulding, W.D., Garbin, C.P., Linszen, D., & Dingemans, P. (1993). Information processing and social cognitive problem solving in schizophrenia: Assessment of interrelationships and changes over time. Journal of Nervous and Mental Disease, 181, 13 - 20.

Perrett, D.I., Hietanen, J.K., Oram, M.W., & Benson, P.J. (1992). Organization and function of cells responsive to faces in the temporal cortex. Philosophical Transactions of the Royal Society of London, Series B, 335, 23 - 30.

Perrett, D.I., & Mistin, A.J. (1990). Perception of facial characteristics by monkeys. In W.C. Stebbins and M.A. Berkley (Eds.), Comparative perception: Volume II: Complex signals (pp. 187 - 215). New York: NY: Wiley.

Perrett, D.I., Mistlin, A.J., Chitty, A.J., Harries, M.H., Newcombe, F., and de Haan, E. (1988). Neuronal mechanisms of face perception and their pathology. In C. Kennard and F.C. Rose (Eds.), Physiological aspects of clinical neuro-ophthalmology (pp. 137 - 154). London: Chapman and Hall.

Phillips, M.L. & David, A.S. (1997). Visual scan paths are abnormal in deluded schizophrenics. Neuropsychologia, 35, 99 - 105

Pilowsky, I. & Bassett, D. (1980). Schizophrenia and the response to facial emotions. Comprehensive Psychiatry, 21, 236 - 244.

Platt, J.R. (1964). Strong inference. Science, 146 (3642), 347 - 353.

Plaut, D.C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. Journal of Clinical and Experimental Neuropsychology, 17, 291 - 321.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired work reading: Computational principles in quasi-regular domains. Psychological Review, 103, 56 - 115.

Popper, K.R. (1965). Conjectures and refutations: The growth of scientific knowledge. New York: Basic books.

Popper, K.R. (1974). The logic of scientific discovery. London: Hutchinson.

Purves, D. (1994). Neural activity and the growth of the brain. Cambridge, UK: Cambridge University Press.

Rakic, P., Bourgeois, J.P., Eckenhoff, M.F., Zecevic, N., & Goldman-Rakic, P.S. (1986). Science, 232, 232 - 235.

Ratcliff, R., van Zandt, T., & McKoon,G. (1999). Connectionist and diffusion models of reaction time. Psychological Review, 106, 261 - 300.

Remington, G., Kapur, S., & Zipursky, R. (1998). The relationship between risperidone plasma levels and domapine $D_2$ occupancy: A positron emission tomography study [CD-ROM]. Journal of Clinical Psychopharmacology, 18, 82 - 83. Abstract from: WinSPIRS: PsychLit Item # 1997 - 38969 - 012. (Review available at http://www.futur.com/resframes.htm.)

Rinehart, M.L. (1995). Learn C++ today. Foster City, CA: IDG Books Worldwide.

Robbins, T.W. (1990). The case for frontostriatal dysfunction in schizophrenia. Schizophrenia Bulletin, 16, 391 - 402.

Rolls, E.T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. Philosophical Transactions of the Royal Society of London, Series B, 335, 11 - 21.

Rossi, E.L. (1996). The symptom path to enlightenment: The new dynamics of self-organization in hypnotherapy: An advanced manual for beginners. Pacific Palisades: CA: Palisades Gateway Publishing.

Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, E.D.Jr., & Beck, L.H. (1956). A continuous performance test of brain damage. Journal of Consulting Psychology, 20, 343 - 350.

Rumelhart, D.E. & McClelland, J.L. (1986). Parallel distributed processing: Explorations in the microstructures of cognition. Cambridge, MA: The MIT Press.

Russell, J.A. (1991). Culture and the categorization of emotions. Psychological Bulletin, 110 (3), 426 - 450.

Russell, J.A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. Psychological Bulletin, 115, 102 - 141.

Salem, J.E., Kring, A.M., & Kerr, S.L. (1996). More evidence for generalized poor performance in facial emotion perception in schizophrenia. Journal of Abnormal Psychology, 105, 480 - 483.

Saugstad, L.F. (1989). Age at puberty and mental illness: Towards a neurodevelopmental aetiology of Kraepelin's endogenous psychoses. British Journal of Psychiatry, 155, 536 - 544.

Scherer, K.R. (1992). What does facial expression express. In K.T. Strongman (Ed.), International review of studies on emotion, Vol. 2 (pp. 139 - 165). New York, NY: Wiley.

Schneider, F., Gur, R.C., Gur, R.E., & Shtasel, D.L. (1995). Emotional processing in schizophrenia: Neurobehavioral probes in relation to psychopathology. Schizophrenia Research, 17, 67 - 75.

Seeman, P. & Lee, T. (1975). Antipsychotic drugs: Direct correlation between clinical potency and presynaptic action on dopamine neurons. Science, 188, 1217 - 1219.

Seeman, P., Lee, T., Chau-Wong, M., & Wong, K. (1976). Antipsychotic drug doses and neuroleptic / dopamine receptors. Nature, 261, 717 - 719.

Seidenberg, M.S. (1993). Connectionist models and cognitive theory. Psychological Science, 4, 228 - 235.

Selemon, L.D., Rajkowska, G., & Goldman-Rakic, P.S. (1995). Abnormally high neuronal density in the schizophrenic cortex. Archives of General Psychiatry, 52, 805 - 818.

Sergent, J. & Signoret, J.L. (1992). Functional and anatomical decomposition of face processing: Evidence from prosopagnosia and PET study of normal subjects. Philosophical Transactions of the Royal Society of London, Series B, 335, 55 - 62.

Siegle, G.J. (1998). Connectionist models of cognitive, affective, brain, and behavioral disorders. World Web Site located at http://www.sci.sdsu.edu/CAL/connectionist-models/.

Silva, J.A., Leong, G.B., & Wine, D.B. (1994). The authors reply. Canadian Journal of Psychiatry, 39, 59.

Silva, J.A., Leong, G.B., Wine, D.B., & Saab, S. (1992). Evolving misidentification syndromes and facial recognition deficits. Canadian Journal of Psychiatry, 37, 574 - 576.

Silverstein, S.M. (1997). Information processing, social cognition, and psychiatric rehabilitation in schizophrenia. Psychiatry, 60, 327 - 340.

Silverstein, S.M. & Schenkel, L.S. (1997). Schizophrenia as a model of context-deficient cortical computation. Behavioral and brain sciences, 20, 696 - 697.

Skinner, H.A. (1975). Model profile analysis: A unified treatment of classification procedures. Unpublished doctoral dissertation. London, ON. University of Western Ontario.

Smith, P.L. (1995). Psychophysically principled models of visual simple reaction time. Psychological Review, 102, 567 - 593.

Soken, N.H. & Pick, A.D. (1999). Infants' perception of dynamic affective expressions: Do infants distinguish specific expressions? Child Development, 70, 1275 - 1282.

Spaulding, W.D. (1997). Cognitive models in a fuller understanding of schizophrenia. Psychiatry, 60, 341 - 346.

Spaulding, W.D., Sullivan, M., Weiler, M., Reed, D., Richardson, C., Storzbach, D. (1994). Changing cognitive functioning in rehabilitation of schizophrenia. Acta Psychiatrica Scandinavica, 90 (Suppl. 384), 116 - 124.

Spiegel, D.E., Gerard, R.M., Grayson, H.M., & Gengerelli, J.A. (1962). Reactions of chronic schizophrenic patients and college students to facial expressions and geometric forms. Journal of Clinical Psychology, 18, 396 - 402.

Spitzer, M. (1997). A cognitive neuroscience view of schizophrenic thought disorder. Schizophrenia Bulletin, 23, 29-50.

SPSS Inc. (1998). SPSS for Windows, Release 9.0.0, Standard Version [computer program]. Chicago, IL: Author.

Stenberg, G., Wiking, S., & Dahl, M. (1998). Judging words at face value: Interference in a word processing task reveals automatic processing of affective facial expressions. Cognition and Emotion, 12, 755 - 782.

Strandburg, R.J., Marsh, J.T., Brown, W.S., Asarnow, R.F., Higa, J., & Guthrie, D. (1994). Continuous-processing related ERPS in schizophrenic and normal children. Biological Psychiatry, 35, 525 - 538.

Suddath, R.L., Casanova, M.R., Goldberg, T.E., Daniel, D.G., Kelsoe, J.R.Jr., & Weinberger, D.R. (1989). Temporal lobe pathology in schizophrenia: A quantitative Magnetic Resonance Imaging study. American Journal of Psychiatry, 146, 464 - 472.

Sullivan, G., Marder, S.R., Liberman, R.P., Donahoe,C.P., & Mintz, J. (1990). Social skills and relapse history in outpatient schizophrenics. Psychiatry, 53, 340 - 345.

Sullivan, E.V., Mathalon, D.H., Zipursky, R.B., Kersteen-Tucker, Z., Knight, R.T., & Pfefferbaum, A. (1993). Factors of the Wisconsin Card Sorting Test as measures of frontal-lobe function in schizophrenia and in chronic alcoholism. Psychiatry Research, 46, 175 - 199.

Symbolic Computation Group (1995). MAPLE V.4 [computer software]. Waterloo, ON: Author.

Taylor, M.A., Berenbaum, S.A., Jampala, V.C., & Clonginger, C.R. (1993). Are schizophrenia and affective disorder related? Preliminary data from a family study. American Journal of Psychiatry, 150, 278 - 285.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. Journal of Mathematical Psychology, 28, 363-400.

Townsend, J.T. & Ashby, F.G. (1983). The stochastic modeling of elementary psychological processes. New York, NY: Cambridge University Press.

Tucker, L.R. & Messick, S. (1963). An individual differences model for multidimensional scaling. Psychometrica, 28, 333 - 367.

Varney, N.R. & Menefee, L. (1993). Psychosocial and executive deficits following closed head injury: Implications for orbital frontal cortex. Journal of Head Trauma Rehabilitation, 8, 32 - 44.

Verdoux, H., van Os, J., Maurice-Tison, S., Gay, B., Salamon, R., & Bourgeois, M. (1998). Is early adulthood a critical developmental stage for psychosis proneness? A survey of delusional ideation in normal subjects. Schizophrenia Research, 29, 247 - 254.

Vickers, D. & Lee, M.D. (2000). Dynamic models of simple judgments: II. Properties of a self-organizing PAGAN (Parallel, Adaptive, Generalized Accumulator Network) model for multi-choice tasks. Nonlinear Dynamics, Psychology, and Life Sciences, 4, 1 - 31.

Waddington, J.T. (1993). Neurodynamics of abnormalities in cerebral metabolism and structure in schizophrenia. Schizophrenia Bulletin, 19, 55 - 69.

Walker, A.S. (1982). Intermodal perception of expressive behaviors by human infants. Journal of Experimental Child Psychology, 33, 514 - 535.

Walker, E. (1994). Developmentally moderated expressions of the neuropathology underlying schizophrenia. Schizophrenia Bulletin, 20, 453-480.

Walker, E., Marwit, S.J., & Emory, E. (1980). A cross-sectional study of emotion recognition in schizophrenics. Journal of Abnormal Psychology, 89, 428 - 236.

Walker, E., McGuire, M., & Bettes, B. (1984). Recognition and identification of facial stimuli by schizophrenics and patients with affective disorders. British Journal of Clinical Psychology, 23, 37 - 44.

Walker-Smith, G.J. (1978). The effects of delay and exposure duration in a face recognition task. Perception and Psychophysics, 24, 63 - 70.

Weinberger, D.R. (1987). Implications of normal brain development for the pathogenesis of schizophrenia. Archive of General Psychiatry, 44, 660 - 669.

Weintraub, S., Mesulam, M.M., & Kramer, L. (1981). Disturbances in prosody: A right-hemisphere contribution to language. Archives of Neurology, 38, 742 - 744.

Where next with psychiatric illness? (1988). Nature, 336, 95 - 96.

White, M. (1999). Representation of facial expressions of emotion. American Journal of Psychology, 112, 371 - 381.

White, M. (2000). Parts and wholes in expression recognition. Cognition and Emotion, 14, 39 - 60.

Wickens, T.D. (1998). On the form of the retention function: Comment on Rubin and Wenzel (1996): A quantitative description of retention. Psychological Review, 105, 379 - 386.

Williamson, P.C. & Conlon, P. (1992). Brian imaging in psychiatry disorders: Present applications and future directions. In E. Persad, S.S. Kazarian, & J.W. Joseph (Eds.), The mental hospital in the 21'st century. Winnipeg, Manitoba: Hignell.

Wölwer, W., Streit, M., Polzer, U., & Gaebel, W. (1996). Facial affect recognition in the course of schizophrenia. European Archives of Psychiatry and Clinical Neuroscience, 246, 165 - 170.

Wong, D.R., Wagner, H.N.Jr., Tune, L.E., Dannals, R.F., Pearlson, G.D., Links, J.M., Tamminga, C.A., Broussoule, E.P., Ravert, H.T., Wilson, A.A., Toung, J.K.T., Malat, J., Williams, J.A., O'Tuama, L.E., Snyder, S.H., Kuhar, M.J., & Gjedde, A. (1986). Positron Emission Tomography reveals elevated $D_2$ receptors in drug naive schizophrenics. Science, 234, 1558 - 1562.

Wyatt, R.J. & Torgorow, I.S. (1976). A comparison of equivalent potencies of neuroleptics as used to treat schizophrenia and affective disorders. Journal of Psychiatric Research, 13, 91 - 98.

Yokota, M., Tokita, G., & Yamamoto, N. (1995). [Facial expression drawn by schizophrenic patients]. Seishin-Igaku [Psychiatric Medicine], 37, 877 - 883.

Young-Browne, G., Rosenfeld, H.M., & Horowitz, F.D. (1977). Infant discrimination of facial expressions. Child Development, 48, 555 - 562.

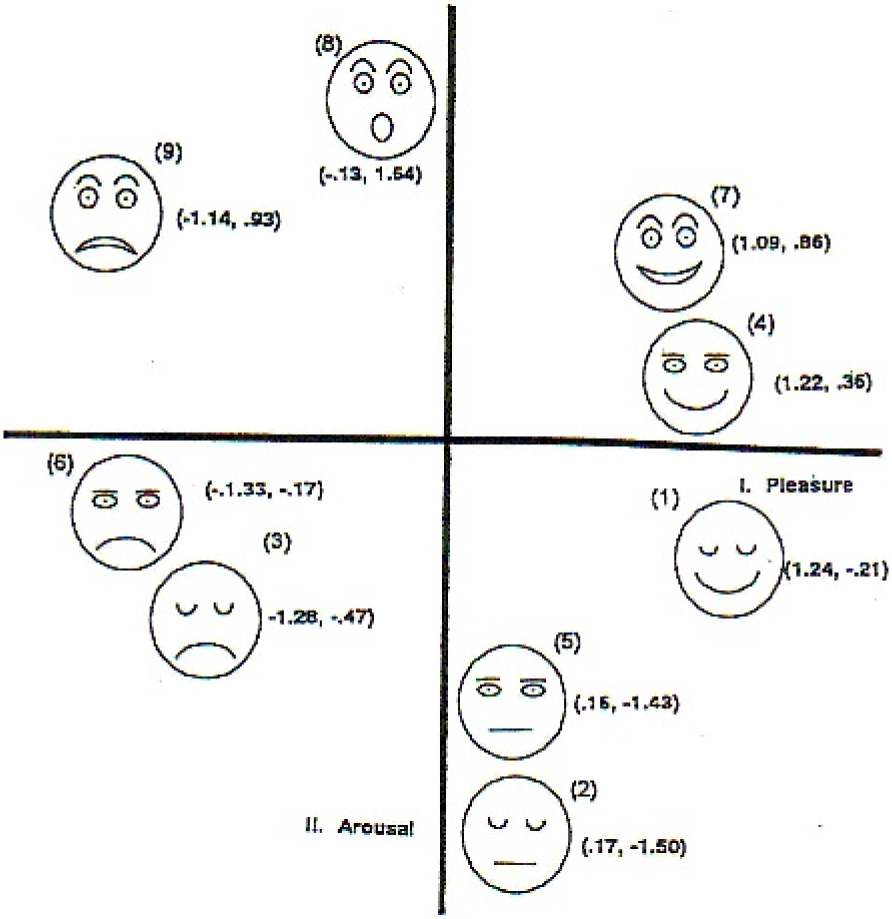
Zec, R.F., Parks, R.W., Gambach, J., & Vicari, S. (1992). The Executive Board system: An innovative approach to cognitive-behavioral rehabilitation in patients with traumatic brain injury. In C.J. Long & L.K. Ross (Eds.), Handbook of head trauma: Acute care to recovery (pp. 219 - 230). New York, NY: Plenum.

Zipursky, R.B., Marsh, L., Lim, K.O., Dement, S., Shear, P.K., Sullivan, E.V., Murphy, G.M., Cseransky, J.G., & Pfefferbaum, A. (1994). Volumetric MRI assessment of temporal lobe structures in schizophrenia. Biological Psychiatry 35, 501 - 516.

Zubin, J. (1972). Scientific models for psychopathology in the 1970's. Seminars in Psychiatry, 4, 283 - 296.

Zucchini, W. (2000). An introduction to model selection. Journal of Mathematical Psychology, 44, 41 - 61.

Zuroff, D.C. & Colussy, S.A. (1986). Emotion recognition in schizophrenic and depressed inpatients. Journal of Clinical Psychology, 42, 411 - 416.



<u>Figure 5</u>. Multi-Dimensional Scaling solution of schematic faces. Taken from Carter & Neufeld, 1996b.

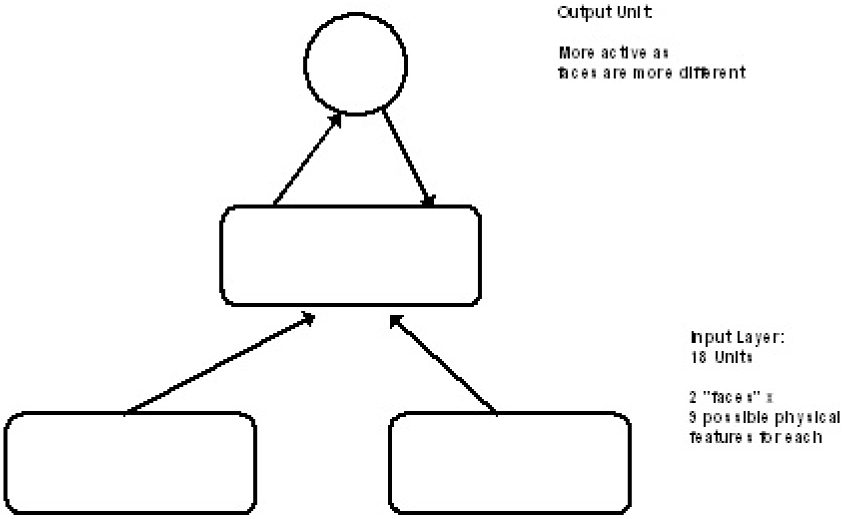
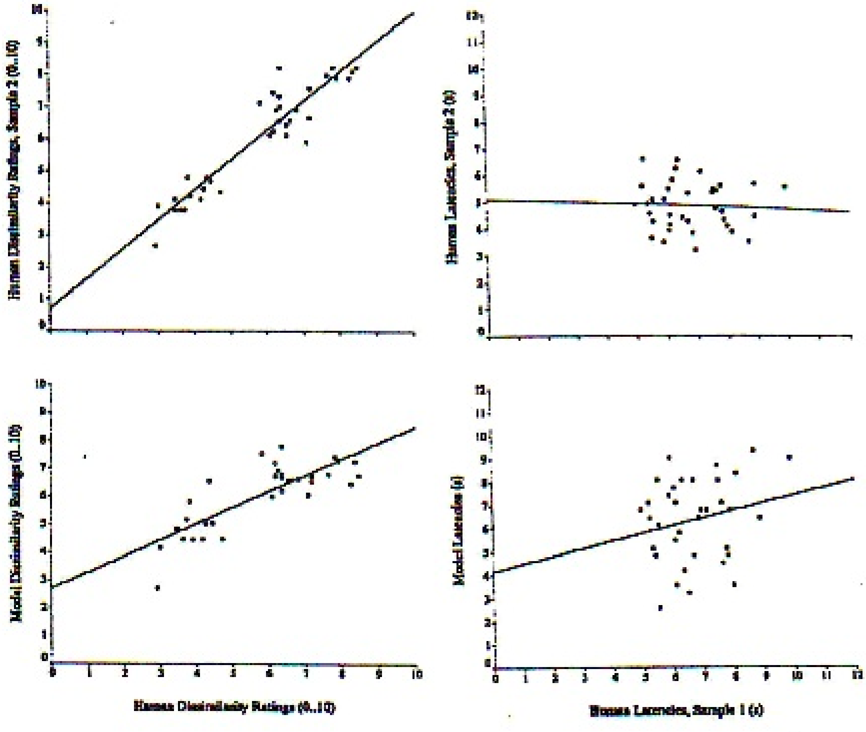


Figure 6: Backpropagation network for schematic faces



<u>Figure 7</u>. Scatterplots of human control group data (top) and model (bottom) in terms of similarity judgments (left, 0..10) and latencies (right, s) for the schematic faces study, with best-fitting regression lines.

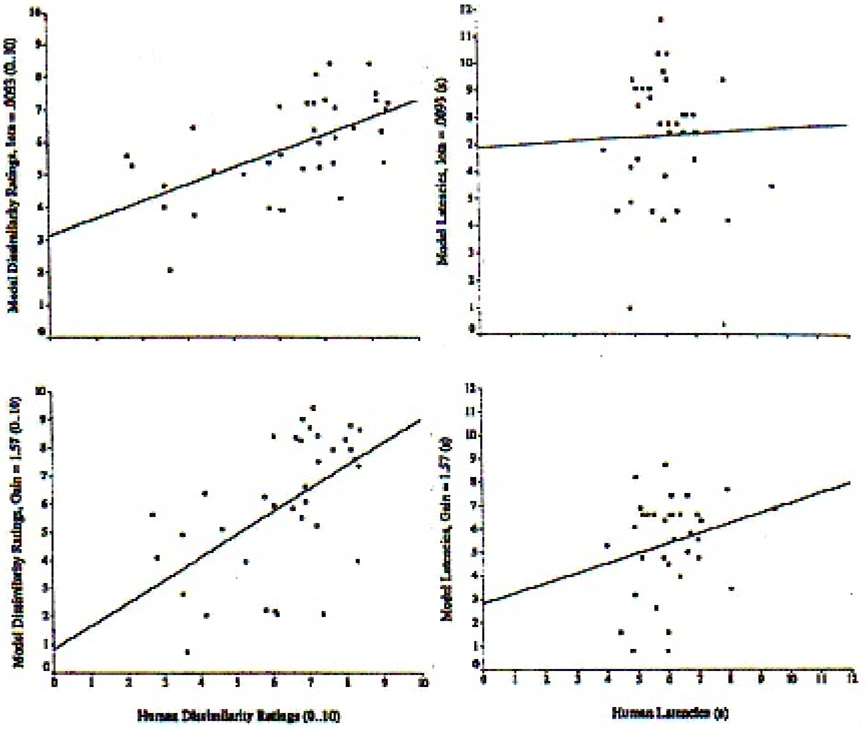


Figure 8. Observed versus predicted similarity ratings (left, 0..10) and latencies (right, s) for the paranoid schizophrenia group, first run of the schematic faces study, $\iota$ = .0093 (upper) and *gain* = 1.57 (lower), with best-fitting regression lines.

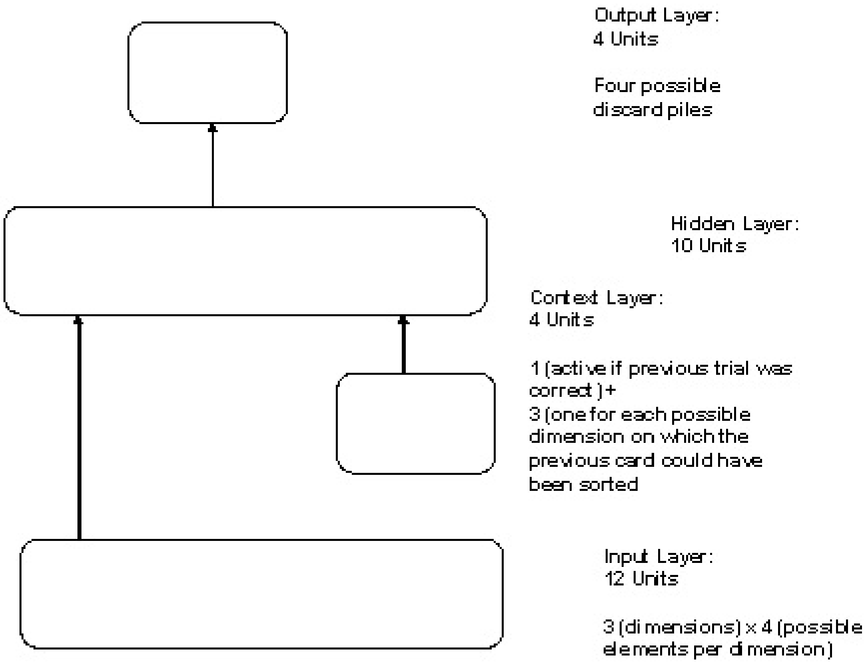


Figure 15: Backpropagation network for the Wisconsin Card Sorting Task

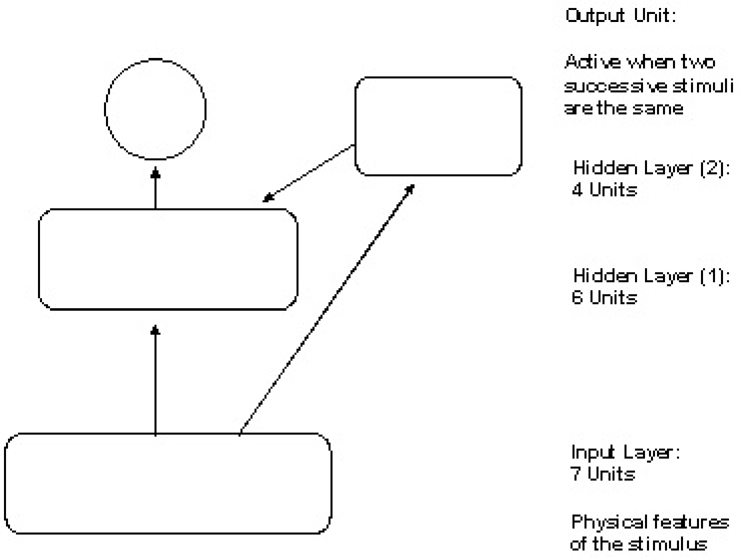


Figure 16: Backpropagation network for the Continuous Performance Task

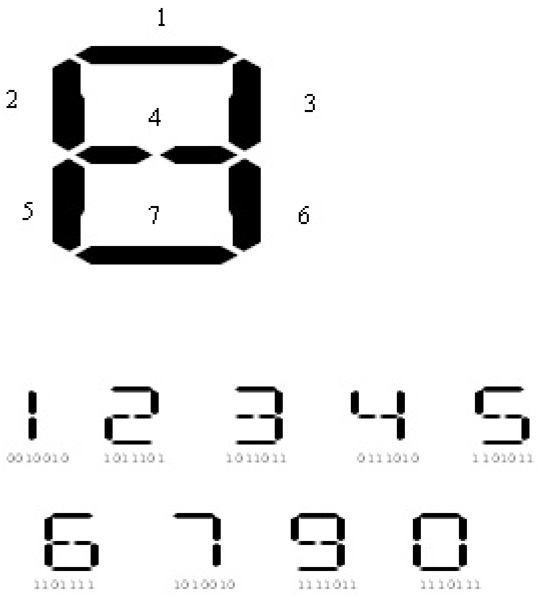


Figure 17: Input for the Continuous Performance Task: Serial positions are indicated at the top, and digits with their binary codes are presented at the bottom.

Table 3

Parameter estimates and empirical fit for alternate cost functions and representative model parameters [second-run results are in square brackets].

| Modeled Hypotheses | | GROUP<br>Paranoid Schizophrenia | | | | Nonparanoid Schizophrenia | | | |
|---|---|---|---|---|---|---|---|---|---|
| Reduced processing stage (sub-process) completion rate; increased *h* | Cost Function | Unmodified Parameter value[a] | Empirical Fit | Parameter Estimate | Empirical Fit[b] | Unmodified Parameter Value | Empirical Fit | Parameter Estimate | Empirical Fit |
| | Provisional $\chi^2$ ("$\chi^2$") | .01373<br>[.04429] | 952.19<br>[863.43] | .02051[c]<br>[.06222] | 354.07[d]<br>[327.56] | .01373<br>[.04429] | 491.38<br>[414.1] | .01636<br>[.04933] | 348.25<br>[333.24] |
| | Sum of Squares *(SS)* | --- | 505.14<br>[349.06] | .02051<br>[.06222] | 350.41<br>[220.55] | — | 229.68<br>[145.04] | .01636<br>[.04933] | 206.39<br>[135.75] |
| | Pseudo $\chi^2$ ($\hat{\chi}^2$) | --- | 1763.2<br>[364.5] | .02051<br>[.06222] | 1,069.96<br>[188.81] | --- | 1094.72<br>[188.45] | .01636<br>[.04933] | 887.013<br>[156.99] |
| Additional sub-processes; decreased $l$ | "$\chi^2$" | .001<br>[.0018] | 952.19<br>[863.43] | .0002<br>[.0003] | 432.64<br>[348.92] | .001<br>[.0018] | 491.38<br>[414.1] | .00045<br>[.00091] | 321.25<br>[318.52] |
| | *SS* | --- | 505.14<br>[349.06] | .00006<br>[.0003] | 120.12<br>[144.24] | --- | 229.68<br>[145.09] | .00045<br>[.00091] | 137.43<br>[111.38] |
| | $\hat{\chi}^2$ | — | 1763.2<br>[364.5] | .00003<br>[.0003] | 42.09<br>[63.04] | --- | 1094.72<br>[188.45] | .00010<br>[.00045] | 64.03<br>[60.65] |

[a] From network trained on control data to a criterion of "$\chi^2_{df}$" *df*-2.

[b] Empirical fit improves as values decrease.

[c] Based on a closed-form function of observed values (see text).

[d] Artefactual deflation of "$\chi^2$" (see text).