Artificial neural networks as analytic tools in an event-related potential study of face memory

Reiko Graham & Michael R.W. Dawson

The Biological Computation Project, University of Alberta, Edmonton, AB, Canada

We compared the abilities of artificial neural networks (ANN's) and ANOVA to classify early latency event-related potentials (ERP's) that were recorded from the right temporal area elicited by recognized and novel faces.

ANOVA was unable to distinguish between the two types of ERP's; however, an ANN was. Network interpretation revealed that classification was achieved through coarse coding in the hidden units. Differences between input time-points that varied according to ERP type were also discovered.

Although further research is needed in order to establish a framework for future analyses, results provide support for the utility of ANN's for ERP analysis and classification.

1. Rationale

Memory can be conceptualized as the result of processing in unimodal and transmodal cortical association areas (Mesulam, 1998). Unimodal areas are modality specific and receive projections from primary sensory cortex. In humans, unimodal visual areas include the fusiform, inferior, and middle temporal areas. Transmodal areas receive inputs from more than one modality and include the prefrontal and posterior parietal cortices.

Neuroimaging (e.g. Kanwisher et al., 1998) and intracranial (e.g. McCarthy, 1997) have identified the fusiform area of the temporal lobe as a unimodal area that is important in face processing. Electrophysiologically, perceptual activity in this area is manifested in the N200 which is maximal over right temporal areas (Bentin et al., 1996).

An important issue is whether activity in this area is also correlated with memory. ERP evidence with humans is mixed. Some studies have reported early latency face memory effects over temporal areas (e.g. Seeck, et al., 1997), while some have not (e.g. Graham & Cabeza, 2001; Muente et al., 1997).

One possibility for this inconsistency is that memory-related voltage changes are represented in ERP's but linear methods of analysis are unable to reliably detect them.

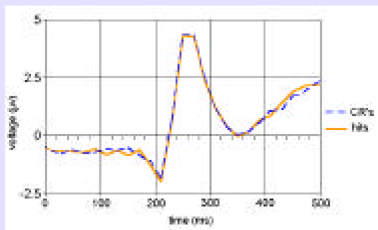
Objectives:

Can ANN's detect differences between early latency ERP's recorded over temporal sites that were elicited by old and new faces?

If an ANN can differentiate between ERP's, how is it doing it? What features of the data appear to be important?

2. Method

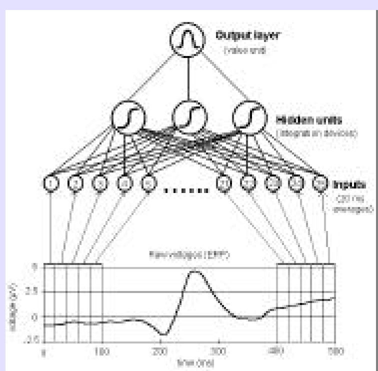
ERP's were obtained for remembered faces (hits) and new faces (correct rejections or CR's) from 42 subjects during a face recognition task. To examine unimodal effects, voltages were taken from the right temporal site (T8). Early ERP's were isolated by taking the first 500ms of the recording epoch. Time-points were averaged into 20ms epochs.



An RM-ANOVA was conducted which included the 25 epochs as predictors and trial type and epoch as within-subjects variables.

Hidden unit 2

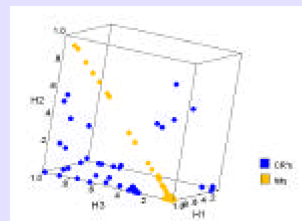
An ANN was trained using the 25 epochs as inputs. We employed a hybrid ANN which used integration devices as hidden units and a value unit as the output unit (integration devices transform the data with a sigmoid function, value units, with a Gaussian function). The ANN had 25 input units, 3 hidden units and 1 output unit.



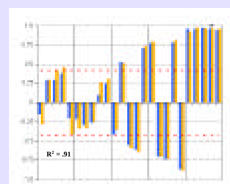
3. Results

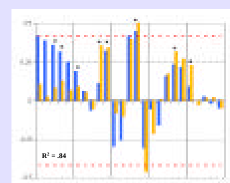
RM-ANOVA did not detect any reliable differences between ERP's to hits and CR's during the first 500ms. The ANN was able to differentiate between the two types.

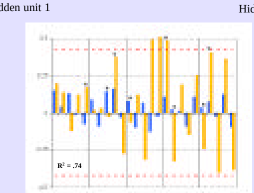
Examination of hidden unit activity revealed that units had relatively non-differential activity to hits and CR's. However, when the three units were examined together, it was possible to see how the discrimination was achieved.



Given that hidden unit activity is a function of net input, we correlated inputs to a hidden unit with its net input, enabling us to determine which time-points had relationships with net input and hence, which may have influenced hidden unit activity.


Hidden unit 1


Hidden unit 2


Hidden unit 3

Correlations revealed relationships which differed depending on ERP type. Step-wise regression confirmed that subsets of inputs accounted for a significant amount of variance in hidden unit activity.

5. Discussion

Early latency information in ERP's from unimodal face processing areas can be used to differentiate between hits and CR's, but this information is in the form of higher-order, non-linear voltage/time relationships.

The ANN can discriminate between ERP's through a form of coarse coding in the hidden units, which may be analogous to encoding in the temporal cortex where faces are represented by patterns of activity in a small number of broadly tuned neurons (Young, 1995).

Examination of the correlations of inputs and net input to a hidden unit revealed differential relationships dependent upon the recognition status of an face.

Much remains to be revealed about how ANN's classify ERP's and a framework for future analyses remains to be established. Nevertheless, preliminary results are encouraging and provide support for unimodal memory effects.

Methods

Subjects: 42 right-handed healthy males and females.
Materials: 240 black and white photographs of unfamiliar male and female faces. Half of the faces were presented during study, and the remainder were used as distractor faces during recognition.
Procedure: Each trial began with the presentation of a fixation rectangle which was replaced by a face for 400 msec. After 1600 msec, subjects were presented with a response selection screen. Trials were separated by one second. During test, subjects indicated which faces were old and which were new
ERP methods: ERP's were recorded from 30 Ag/AgCl electrodes. EEG was sampled for an epoch of 1700 msec, starting 100 msec prior to the onset of a face.

References

Bentin, S. et al. (1996). Jnl of Cog Neurosci, 8(6), 551-565.
Graham, R. & Cabeza, R. (2001). Neuroreport, 12(2), 245-248.
Kanwisher, N. et al. (1998). Cognition, 68, B1-B11.
McCarthy, G.et al. (1997). Jnl of Cog Neurosci, 9(5), 605-610.
Mesulam, M.-M. (1998). Brain, 121, 1013-1052.
Muente, T.F. et al. (1997). Neuroscience Research, 28, 223-233.
Seeck, M. et al.(1997). Neuroreport, 8, 2749-2754.
Young, M.P. (1995). In M. Gazzaniga (Ed.), The cognitive neurosciences (pp. 463-474) . Cambridge, MA: MIT Press.