A connectionist model of categorization response times

David Peebles
University of Nottingham

Koen Lamberts
University of Warwick

Abstract

Although perceptual categorization has been studied extensively in psychology, response times in categorization tasks have only recently become an important research topic [1, 2, 3, 4, 5, 6, 7, 8, 9]. In this article, we propose a connectionist model of categorization RT, called CONCAT, which aims to provide a joint account of response times and choice proportions in binary classification tasks. First, we outline the basic principles of the model. Next, we present a perceptual categorization experiment and apply CONCAT to the results.

1 Principles of CONCAT

CONCAT is based on the EGCM-RT, which is a formal model of categorization RT proposed by [5]. The EGCM-RT is not implemented as a connectionist model. CONCAT is a fairly direct translation of the EGCM-RT into a connectionist framework. In this chapter, we cannot discuss all aspects of the EGCM-RT that are relevant in the current context (see [5] for an in-depth discussion of the EGCM-RT). Instead, we will describe the basic principles of CONCAT without direct reference to its precursor, and without disussing the extensive evidence that has been collected in support of the EGCM-RT (which also supports CONCAT). The main difference between CONCAT and the EGCM-RT is that the EGCM-RT has a broader scope than CONCAT. CONCAT applies only to stimuli with separable and binary dimensions, whereas the EGCM-RT also explains RTs with stimuli that have continuous and/or integral dimensions. A connectionist implementation of the full EGCM-RT will be reported elsewhere.

The architecture of CONCAT is presented in Figure 1. The model takes the form of a 4-layer radial-basis function network (see [2]). The input nodes encode the stimulus. There is one input node for each stimulus dimension. The state of the input nodes is characterized by two numerical parameters. The first parameter (x_{ip}) is the value of stimulus i on dimension p represented by the node. In the case of binary stimuli, 0 and 1 are used as possible dimension values. The second parameter (φ_p) refers to the activation state of the node. Input nodes are either active ($\varphi_p = 1$) or not ($\varphi_p = 0$).

Stimulus encoding is not all-or-none. Instead, the encoding process involves the activation of input nodes over time, which corresponds to a process of sampling stimulus information. The probability that input node x has been activated at or before time t after stimulus onset (called the cumulative inclusion probability, see [3, 4, 5]) is given by an exponential distribution function:

$$i_x(t) = 1 - \exp(-q_x t) \tag{1}$$

In this equation, q_x is the inclusion rate (or processing rate) of dimension x. Input nodes that correspond to dimensions with a higher processing rate are usually activated faster than nodes that correspond to dimensions with low processing rates.

The exemplar units correspond to individual stimuli that were stored during category learning. Whenever an input unit is activated, the activation of the exemplar nodes is updated.

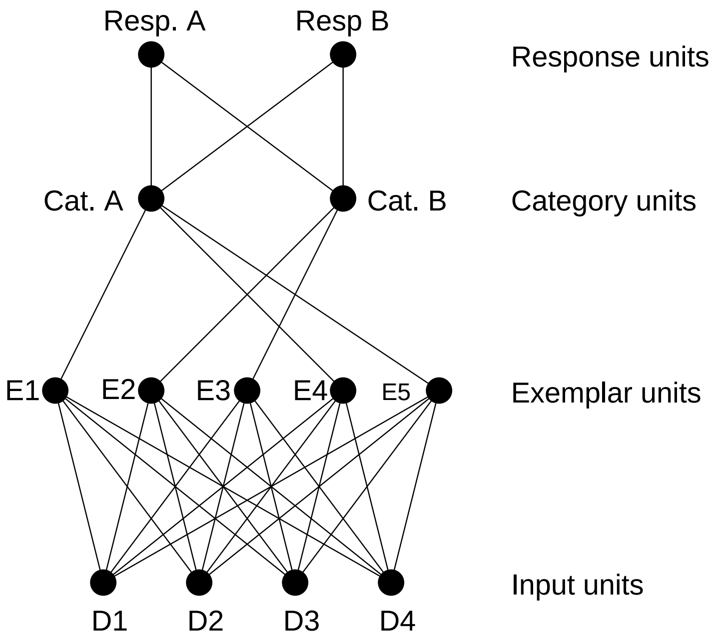

Figure 1: Architecture of CONCAT. A full description of the model's components is provided in the main text.

Exemplar node activation equals the similarity between the current input pattern and the pattern that corresponds to the stimulus they encode. Similarity is defined as

$$s_{ij}\left(\Phi\right) = \exp\left[-c\left(\sum_{p=1}^{P} u_p\left[\varphi_p\left|x_{ip} - x_{jp}\right|\right]^r\right)^{q/r}\right] \tag{2}$$

in which $s_{ij}(\Phi)$ is the similarity between stimulus i and stored exemplar j given set Φ of activated input nodes, c is a generalization value, p is an index for the dimensions (the total number of dimensions is P), u_p is the utility value of dimension p ($0 \leq u \leq 1, \sum u = 1$), φ_p indicates whether the node for dimension p is active, and x_{ip} and x_{jp} are the values of the stimulus and the stored exemplar on dimension p. In all applications in this article, q was set equal to 1 (see [5]). The utility value of a dimension indicates how important that dimension is in the similarity computation. Dimensions that are highly diagnostic tend to have a high utility value (see [3, 7]). As a consequence of the time course of input node activation and the similarity computation that takes place in the exemplar nodes, the activation of the exemplar nodes changes over time, as more stimulus dimensions are processed.

In the third layer of the network, the category units compute the total similarity of the input pattern to all exemplars from each category. This is achieved by letting the activation of each category node correspond exactly to the sum of the activations of the exemplar nodes that belong to the category.

Finally, whenever an input unit has been activated and activation has spread to the category nodes, a decision is made as to whether sufficient stimulus information has been acquired to stop sampling and initiate a response, or whether more stimulus information is needed. This decision is made at the highest level of the network, that of the response units. The summed similarity of the stimulus to all exemplars from one category is divided by the summed similarity of the stimulus to all relevant exemplars in memory. If the total similarity to A exemplars is high relative to the total similarity to B exemplars, the probability that the subject will stop sampling and produce an A response is relatively high compared to the probability of stopping sampling and producing a B response. The probability that sampling will stop (for stimulus i and set of active input units Φ) and that a category A response will be given equals

$$P\left(Stop\ \&\ A|i,\Phi\right)=\left[\frac{b_A S_{iA}\left(\Phi\right)}{b_A S_{iA}\left(\Phi\right)+\left(1-b_A\right)S_{iB}\left(\Phi\right)}\right]^{\theta}. \tag{3}$$

In this equation, b_A is the response bias for category A, $(0 \leq b_A \leq 1)$, and S_{iA} is the activation of category node A. The parameter θ can have any value larger than or equal to 1. This choice ratio can also be interpreted as a measure of *confidence* in the category membership of the stimulus. If the subject is very confident about category membership, there is a high probability that sampling will stop and a response will be initiated.

Analogously, the probability that sampling will stop and a category B response will be given is

$$P\left(Stop\ \&\ B|i,\Phi\right)=\left[\frac{b_B S_{iB}\left(\Phi\right)}{\left(1-b_B\right)S_{iA}\left(\Phi\right)+b_B S_{iB}\left(\Phi\right)}\right]^{\theta}. \tag{4}$$

If all stimulus dimensions have been sampled, the stopping probability is defined as 1 and a response will be initiated immediately. A further assumption is that

$$P\left(Stop|i,\oslash\right)=0, \tag{5}$$

which means that at least one dimension will be sampled before stopping.

To summarize, CONCAT assumes that (i) stimulus dimensions are sampled stochastically in the earliest stages of categorization, and (ii) that the relative summed similarity to exemplars from the alternative categories determines the probability that sampling is interrupted and a response is initiated. The model predicts response time differences between stimuli in terms of the expected duration of dimension processing, and in terms of the number of dimensions that are processed on individual trials.

2 Experimental test of CONCAT

The primary purpose of this experiment is to test the ability of CONCAT to account for joint categorisation accuracy and RT data from a standard speeded category learning experiment.

2.1 Method

2.1.1 Participants

Ten undergraduate and postgraduate psychology students from the University of Birmingham participated in the experiment [9]. The undergraduates who took part were given credit towards the Psychology department's research participation scheme.

2.1.2 Apparatus and stimuli

The experiment was carried out on an Elonex PC-433 computer with a Vale EC 33 cm SVGA colour monitor using a display mode with 640 pixels horizontally and 480 pixels vertically. Participant's responses were registered by two microswitches connected to the computer's parallel port. The stimuli used were drawings of aeroplanes viewed from above which varied on four binary dimensions—shape of nose (round or pointed), shape of wings (straight or tapered), number of engines (two or four), and shape of tail (square or rounded). Two example stimuli showing the full range of dimension values are shown in Figure 2.

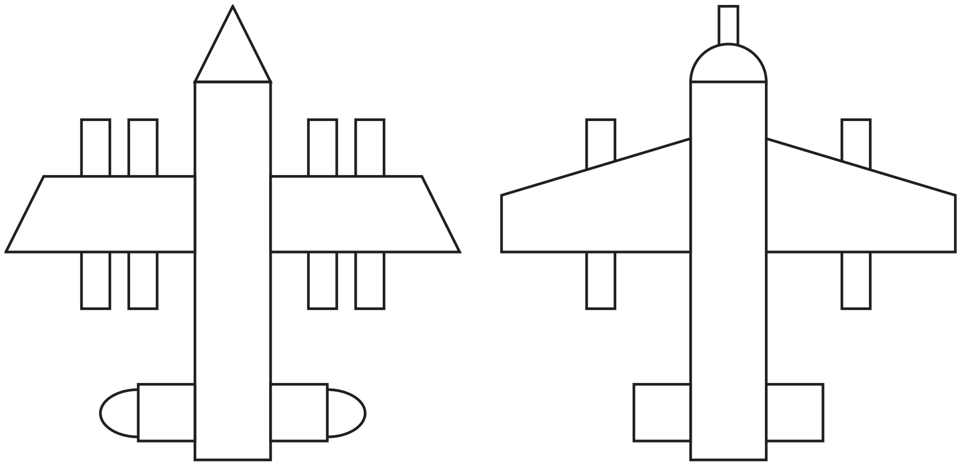

Figure 2: Sample stimuli used in the experiment

2.1.3 Design and procedure

The category structure, which comprised a total of eight stimuli, is shown in Table 1. The structure is regular in that each stimulus differs on two dimensions from six of the other stimuli and on all dimensions from one stimulus in the alternative category. This regularity is also apparent in the arrangement of the stimuli as on each dimension, three of the four stimuli in a category have one value while one stimulus has the opposite value. In addition, no dimension is more predictive of a category than any other.

The experiment was a standard categorisation reaction time experiment which consisted of a training stage, in which participants were required to learn to classify stimuli into two categories, followed by a transfer stage, where the task was to classify the same stimuli again as quickly as possible without sacrificing accuracy. In both training and transfer stages, participant's category responses were recorded and in the transfer stage, the time of each response (in milliseconds) was also recorded. In the training stage, participants were presented with the stimuli in blocks, each block consisting of the complete set of training stimuli presented sequentially in random order. Training continued until two blocks in succession had been categorised correctly. On each training trial, a white fixation cross would appear at the centre of the blank computer screen for a period of 400 ms followed by a period of 100 ms where the screen was again blank. Then one stimulus chosen at random would appear at the centre of the screen. When one of the two response buttons was pressed, an auditory signal indicating the correctness of the response was given for a period of 500 ms and the screen would be cleared. If the category response was correct, a 600 Hz (high) tone was given whereas if it was incorrect, a 100 Hz (low) tone was given. An interval of 1500 ms separated each training trial. To eliminate any effect of response bias due to hand preference, category labels were randomly assigned to left and right response buttons.

After a short break, participants underwent a transfer stage in which blocks of stimuli were presented again as during the training stage. Trials in the transfer phase were identical to training trials except that no auditory feedback was given. Participants were instructed to categorise the planes as before but this time to be as fast as they could while trying to remain as accurate as possible. Each participant categorised a total of fifty blocks in the transfer stage, being allowed to rest for a few minutes twice during the session, after the completion of 17 and 34 blocks.

Table 1: Structure of stimuli used in the experiment

Structure	Stimulus Number	Stimulus Dimension			
		Nose	Wings	Engines	Tail
Category A	1	1	1	1	0
	2	1	1	0	1
	3	1	0	1	1
	4	0	1	1	1
Category B	5	0	0	0	1
	6	0	0	1	0
	7	0	1	0	0
	8	1	0	0	0

2.2 Results

2.2.1 Training

The mean number of blocks required to achieve two consecutive correct blocks was 42.2 ($SD =$ 23.1). A simple measure of the difficulty of learning the category structure is the mean error frequency for each stimulus over the course of training across participants. These frequencies are presented in Table 2.

Table 2: Mean error frequencies during training

Stimulus	Error Frequency
1 1110 (A)	11.2
2 1101 (A)	21.4
3 1011 (A)	12.3
4 0111 (A)	12.3
5 0001 (B)	8.50
6 0010 (B)	14.8
7 0100 (B)	13.3
8 1000 (B)	12.9

An analysis of variance (ANOVA) on the mean correct responses yielded a significant effect of stimulus $F(7, 63) = 2.49$, $p < .05$, $MSE = .062$. In the training stage, participants' errors were generally evenly distributed across the stimuli, with the exception of stimulus 2 which was on average misclassified more often and stimulus 5 which was in general categorised more accurately. The fact that stimuli 2 and 6 had the highest error rates suggests the possibility that participants were paying greatest attention to the engines dimension as these two stimuli differed from the others in their category on this dimension. Conversely, the low error rates of stimuli 1 and 5 suggests that relatively little attention was being paid to the tail dimension during the training stage.

According to an MDS analysis of similarity, if attention is focused upon a particular dimension, the psychological space in which the stimuli are represented is stretched along that dimension, resulting in a decrease in similarity between stimuli which differ on the dimension. Conversely, if relatively little attention is paid to a dimension, perceived differences between stimuli on that dimension will be reduced, resulting in an increase in similarity between stimuli which differ on the dimension. In terms of the data presented in Table 2, a relatively high level of attention to the engines dimension will decrease the similarity between stimulus 2 and the category A stimuli and stimulus 6 and the category B stimuli. Therefore, the probability that they are classified correctly as belonging to their respective categories will be reduced. In contrast, according to the same analysis, the relatively high levels of classification accuracy for stimuli 1 and 5 suggests that perceived differences that these stimuli have with other stimuli in their respective categories have been minimised, leading to the conclusion that the dimension upon which they differ is receiving relatively little attention.

Given that the category structure is such that no dimension is more predictive than any other, one plausible explanation for the uneven distribution of participants' attention is that stimulus dimensions may be more or less salient than others, (i.e. the engines dimension is highly salient whereas the tail dimension is not particularly salient). It is important to note, however, that effects of differences in salience, reflected by mean error rates from the entire course of training, can be expected to reduce as training progresses as participants are required to attend to all dimensions equally in order to achieve the criterion level of accuracy.

2.2.2 Transfer

The proportions of category A responses and mean RTs for each stimulus are shown in Table 3. An ANOVA on the mean RTs produced a significant effect of stimulus $F(7,63) = 6.81$, $p < .001$, $MSE = 73,617$. An ANOVA on the proportions of correct responses found no significant effect of stimulus. This latter result is possibly due to the fact that participants had been trained to a relatively high criterion of performance in the training stage, as is evidenced by the high levels of accuracy for all stimuli in the transfer stage (mean correct response proportion over all stimuli = .91, $SD = 0.022$).

Table 3: Proportions of category A responses (RP) and mean response times (RT in milliseconds) for each stimulus.

Stimulus	RP	RT
1 1110 (A)	0.904	1061
2 1101 (A)	0.878	1125
3 1011 (A)	0.916	1104
4 0111 (A)	0.932	1041
5 0001 (B)	0.070	854
6 0010 (B)	0.100	1077
7 0100 (B)	0.074	981
8 1000 (B)	0.124	1058

To test the hypothesis that more errors during category learning are accompanied by slower RTs during transfer, the mean error frequencies from the training phase were correlated with the mean RTs. The resulting correlation coefficient was .68. This moderately high positive correlation can be confirmed by studying the values for individual stimuli. Stimulus 5, for example, has the lowest error frequency in the training stage and also has the shortest RT and one of the highest accuracy rates in the transfer stage. The opposite pattern is found with Stimulus 2, which has the highest error frequency in the training stage and the longest RT and one of the lowest accuracy rates in the transfer stage.

The error rates in the transfer stage were also correlated with mean RTs, producing a correlation coefficient of .61. Again, this correlation is supported by the values of individual stimuli. For example, the three stimuli with the highest accuracy rates in the transfer stage—stimuli 4, 5 and 7, are the stimuli with the shortest mean RTs and the stimulus with the lowest level of accuracy-stimulus 2, has the longest mean RT. There are several differences between accuracy levels in the transfer stage and error rates in the training stage (the correlation between these sets of values = -.64). For example, stimulus 4 is classified most accurately in the transfer stage although the error frequency for that stimulus in the training stage is not particularly low. In addition, stimulus 7 is also accurately classified in the transfer stage despite having a relatively high error frequency in the training stage. It should be remembered, however, that the differences in the accuracy levels between the stimuli in the transfer stage are very small.

3 Model-based analysis

Before applying CONCAT to the data, one may attempt to anticipate some of the parameter values it will estimate based upon knowledge of the model's processing assumptions and the observed data. For example, the low error frequency in the training stage and high level of accuracy and short RT in the transfer stage for stimulus 5 may suggest that the low salience of the tail dimension continued to affect categorisation performance throughout the entire experiment. Similarly, the relatively low accuracy levels and long RTs for stimuli 2 and 6 may suggest that the engines dimension was particularly salient and that this also affected performance in the transfer stage, despite the high training criterion. One may expect CONCAT to allocate low and high values for the inclusion rate parameters to these dimensions respectively. The high level of accuracy for stimulus 4 can also be expected to result in high attention and inclusion rate values for the nose dimension.

CONCAT was applied jointly to the category A response proportions and RT data. The predicted category A response proportions and RTs produced by the model are displayed in Table 4. Best fitting parameter values were found by using a search algorithm that maximised the summed coefficient of variation (R^2) for category A response proportions and RTs. Category A response proportions were used for model optimisation rather than proportions of correct responses in order to maximise the variability in the response proportions and so reduce their effect in the estimation of total goodness-of-fit. The primary reason for doing this is to increase the effect of RT data on model optimisation because the ability of the model to predict RTs is the main focus of this investigation.

Table 4: Observed (Obs) and predicted (Pred) category A response proportions (RP) and response times (RT in milliseconds) for each stimulus.

	RP		RT	
Stimulus	Obs	Pred	Obs	Pred
1 1110 (A)	0.904	0.981	1061	1062
2 1101 (A)	0.878	0.858	1125	1096
3 1011 (A)	0.916	0.935	1104	1102
4 0111 (A)	0.932	0.950	1041	1091
5 0001 (B)	0.070	0.001	854	860
6 0010 (B)	0.100	0.001	1077	1061
7 0100 (B)	0.074	0.001	981	976
8 1000 (B)	0.124	0.001	1058	1050

CONCAT had eleven free parameters: four dimension processing rates q, a generalisation parameter c, three dimension utility values u, (the utility value of the fourth dimension is constrained by the values of the other three as all utility values are required to sum to 1), a category response bias parameter b, the parameter θ in the power function which determines the expected duration of dimension processing and a residual time parameter t_{res}. The best fitting parameter values estimated for the model are shown in Table 5.

The model provided a good fit to both the RT data ($R^2 = .93$) and a reasonable fit to the choice proportion data ($R^2 = .98$). In particular, the model was able to predict the short RTs for stimuli 5 and 7 and also predicted that stimuli 2 and 3 had the longest RTs. As expected, the dimension inclusion rate parameter values estimated by the model indicate that the nose and engines dimensions were most salient and that the tail dimension was the least salient of all. The utility parameter for the tail dimension also had the lowest value (0.000), indicating that this dimension was not taken into account at all. The category bias parameter was relatively low, indicating that, according to the model, participants had a slight tendency to favour a category B response (note that a value of $b = 0.5$ indicates no category bias, $b > 0.5$ indicates a category A bias and $b < 0.5$ indicates a category B bias). Because of the relationship between response accuracy and RT embodied by the model, this is probably due to the fact that the average observed RT for category B stimuli is 360 ms shorter than that for the category A stimuli. This is also reflected in the very low response proportions for the category B stimuli predicted by the model.

Table 5: Best-fitting parameter values for CONCAT. Note: The value of the utility parameter for the tail dimension (in brackets) is constrained by the utility values of the other three.

Parameter	Value
$q(nose)$	97.564
$q(wings)$	0.129
$q(engines)$	37.583
$q(tail)$	0.003
θ	11.871
t_{res}(ms)	723
$u(nose)$	0.298
$u(wings)$	0.342
$u(engines)$	0.360
$[u(tail)]$	0.000
c	4.284
b	0.345

4 Discussion

The main reason for conducting this experiment was to test the CONCAT model on data from a standard speeded category learning experiment and to analyse its predictions. CONCAT provided a close fit to the data. The flexibility of CONCAT in terms of its ability to accommodate dimensional salience and attention distribution separately is likely to be the main reason for the model's close fit to the combined data.

The success of the model in accounting for the data provides further support for the assumption that RT differences between stimuli can be explained in terms of the time course of feature processing (see [5]). Further research is needed to determine whether the time course of feature processing is sufficient to explain RT differences between stimuli in a wider range of categorization tasks, or whether other mechanisms are needed as well (e.g. [8]).

References

[1] F. G. Ashby, G. Boynton, and W. W. Lee. Categorisation response time with multidimensional stimuli. *Perception & Psychophysics*, 55:11–27, 1994.

[2] J. K. Kruschke. ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99:22–44, 1992.

[3] K. Lamberts. Categorisation under time pressure. *Journal of Experimental Psychology: General*, 124:161–180, 1995.

[4] K. Lamberts. Feature-accumulation theory of speeded categorization. *Psychological Review (in press)*, 1998.

[5] K. Lamberts. The time course of categorisation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24:695–711, 1998.

[6] K. Lamberts and R. P. J. Freeman. Building object representations from parts: Tests of a stochastic sampling model. *Journal of Experimental Psychology: Human Perception and Performance (in press)*, 1998.

[7] R. M. Nosofsky. Attention, similarity, and the identification-categorisation relationship. *Journal of Experimental Psychology: General*, 115:39–57, 1986.

[8] R. M. Nosofsky and T. J Palmeri. An exemplar-based random walk model of speeded classification. *Psychological Review*, 104:266–300, 1997.

[9] D. Peebles. *The effect of stimulus frequency on classification accuracy and response time*. PhD thesis, University of Birmingham, Birmingham, 1997.

Published in Heinke, D., Humphreys, G. W., & Olson, A. (Eds.), (1999). *Connectionist Models in Cognitive Neuroscience*, pp. 228–239, London, Springer-Verlag.