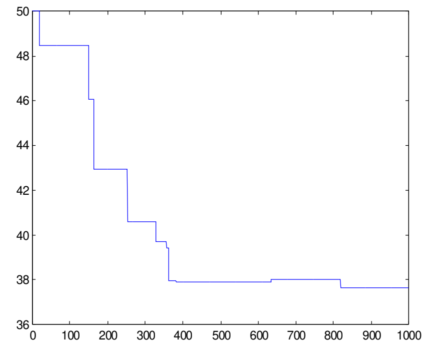
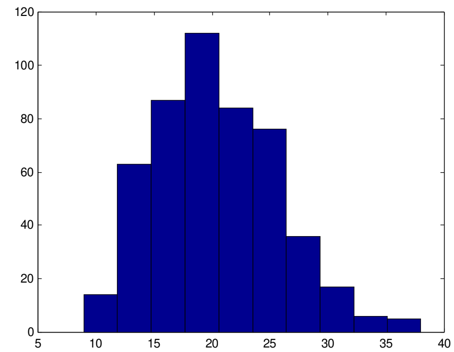
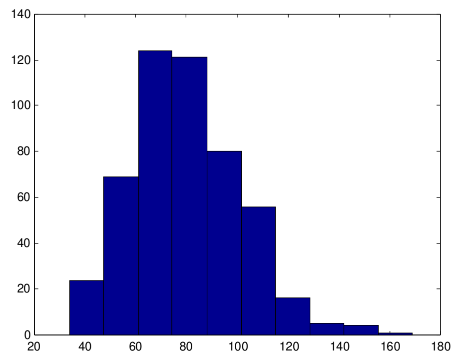
Figures 9.1–9.4: Data form the bisection timing network







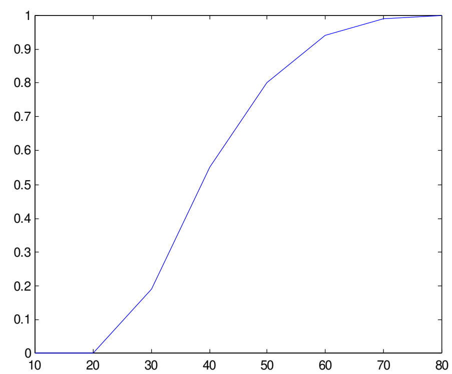


Figure 9.1 displays the normalised output value Sm as a function of alternating training trials of S (40 time steps) and L (80 time steps). Figure 9.2 and 9.3 displays the normalised distribution of accumulator values for 40 and 80 time steps from 500 simulation trials. Figure 9.4 shows the probability of an L response as a function of interval length. As can be observed by eye, the point where 50% of responses were made was approximately 38 time steps, the value of the learnt discriminant (Figure 9.1).

9.6 Discussion

For the simulations reported here, the best discriminant depends on the difference in the SD of accumulator distributions as well as the ratio of the two times. We shall call the proposal that the animal learns the most optimal discriminant the ‘ bisection discrimination hypothesis’ (BDH). Another view is that bisection performance may or may not depend on discrimination but still depends on the SD differences between accumulator distributions S and L. We shall call this the ‘bisection variance hypotheses’ (BVH). If the BVH is true, this would not imply that the BDH is also true.

One way to test the BDH is to investigate the following. Imagine that a bisection experiment is performed with test intervals above L and below S as well as the normal intermediate intervals. If animals learn explicit representations of S and L the probability of a response for a range of intervals above and below L (or S) might be expected to yield a close to symmetrical response distribution as observed in PI performance. The absence of this pattern would not necessary disprove that animals have two explicit representations because the bisection task involves classification rather than reproduction. It is possible and certainly plausible that animals use different cognitive processes for the two different tasks. However, if animals did show close to symmetrical response distributions for S and L, this would violate the view that they only learn a discriminant.

A way of testing the ‘BVH’ is to vary the interval S in training. This rests on the assumption that under certain conditions the animal cannot determine whether the variability of a presentation of S or L is due to the real variability of the stimuli or error in their own accumulator representation. Clearly, there must be some degree of variability in the presentation of an interval stimulus in which the animal can discern that it is the stimulus rather than its own cognitive systems. We shall assume that this degree of variability is quite large. If the BVH was true then one would predict that if the cognitive system experienced the two accumulator distributions with the same variability, arithmetic mean bisection would be predicted. In training if one presented S with a mean of S and some SD this would increase the SD of the experienced accumulator values for the interval S. Under such conditions one would predict that the bisection point would be shifted closer to the arithmetic mean of S and L than when this manipulation was not employed.

In this model the idea of a relative response thresholds was not employed. This did not seem a sensible thing to do as responses are based on deciding whether Sa was above or below the discriminant value. Note that if a process was implemented that involved computing differences between the discriminant and accumulator value relative to the discriminant value, there would not be an affect of relative calculations. Note also that where differences are computed relative to explicit times of S and L (I.e. abs (S−a)/S and abs (L−a)/L), and responses are made for the lowest value on each stimulus presentation, geometrical mean bisection is not obtained. Rather some value less than the geometrical mean is predicted (with S at 2 and L at 8 bisection occurs at 3.6) and was observed in computer simulations performed by the author using the accumulator neural network.

As discussed when the CV value was above .26 the discriminant was moved away from the geometrical mean closer to the value of S. The average CV value for rats and pigeons is usually in the range of .2 to .3 although, this depends on the way CV is calculated.[1] In addition, although the ratios used are commonly 1:4 for S and L, accumulator distributions that overlap for some CV value at some interval ratio will not necessarily overlap when the ratio is increased. For example the accumulator distributions for S and L with the 1:4 ratio that only just overlapped when CV was .26 did not overlap at a 1:8 ratio (see appendix 1). This limitation is in line with the observation that near geometrical mean bisection seems only to occur with ratios not greater than 1:4 (e.g. Siegal 1986).

If one used the learning technique reported here, in cases where the distributions do not overlap, as with larger ratios, the discriminant could settle at a number of different

1 This information came from John Wearden who is actively involved in timing research.

locations. Exactly where it would settle depends on the starting value of the discriminant. If the weights were initialised with a Sm value between values of the two distributions, it would not move, if below the maximum S accumulator value it would settle just above the maximum S value, if above the minimum L value it would settle just below the minimum L value. This would predict rather arbitrary bisection points for large ratios.

The question of whether BDH is worthy of further development appears negative in view of the observation that animals do not always express a bisection point under certain experimental conditions. [2] In studies where large ratios were used, animals were observed to show periods of not responding for a range of intermediate values between S and L (e.g. Platt and Davis, 1983). This is suggestive that animals do not respond L when the stimulus is sufficiently unlike S. More importantly, this is not in line with the discrimination model reported here, in which responses are made for L when above the discriminate and S when below the discriminate. Instead, the data is in line with the view that the animals learn explicit representations of S and L.

A last point worthy of consideration concerns the work of Rodriquez–Girones and Kacelnick (1998). These researchers looked at response latencies in bisection tasks performed by human subjects. Amongst other findings they observed that response latencies were greater when the stimuli was close to the geometrical mean of the standards. The authors conceptualised there findings in the framework of SET in which longer latencies are believed to reflect a subject having to take more samples from a reference memory to find a value that was close to the interval stimuli. Whilst this explanation has very little plausibility the research findings are of interest in conceptualising bisection processes. In the hypothetical case that the animal does in fact learn two explicit times (as data suggests), a greater delay observed at intermediate

2 This research only came to the author's attention at a very late stage in this research following the computer simulations.

times suggests that some form of increased response competition occurs. The greater the competition between representations the harder to resolve and the longer this process would take. The competition of representations suggest processes more sophisticated than modelled here and is a notion congruent with connectionist research, particularly 'self organising' architectures employing some kind of 'rich get richer effect' by recurrent feedback. These kinds of recognition processes operate by competition. Each representation inhibits the other. The representation with the highest excitation has the greatest inhibitory affect on other representations. This permits the winning representation to gradually become enhanced whilst the other competing representations become attenuated. It is suggested that research on such connectionist models may be of benefit to developing models of temporal bisection.

SECTION 10

General Discussion

10.1 Comparison with other models

One important issue concerns how the models reported here compare with models discussed in section 6. The accumulator model reported here improves on the accumulator model proposed by Miall (1993) because it provides a representation of elapsed time with a mean and a SD that increases linearly. Miall describes his accumulator model with less abstraction than the description than the model reported here. Miall's description, however, provides no explanation of how or why neurons have a small probability of being activated or deactivated that allows the gradual accumulation of activity. As a result, Miall's model is no less abstract than this model because activation units can be viewed as single neurons that become active and remain active for only a single time step.

Like Church and Broadbent's oscillator model and Miall's accumulator model, the models reported here provide a quantifiable representation of memory and immediate *elapsed time. The interval based timing schemes such as the spectral timing model (Grossberg and Merril 1992, 1998) and some of the time dependent spatial pattern models (Buonomano and Mauck, 1994; Buonomano and Merzenich, 1995) do not provide such a representation. The question of whether models that don't comprise of quantifiable representations could explain tasks like time−left shall be discussed below.*

As abstract as the models reported here are, they include an attempt to consider timing processes not included in the standard model. Firstly there is the idea that an expectancy representation is formed and persistence of this representation in timing tasks reflect sustained attention or a working memory of a reward anticipated event. An important benefit of adopting this view is that having a sustained expectancy representation or processes that are involved in sustained attention, can also explain the persistence of accumulation and continued comparisons made by the cognitive system. Secondly there is the view that timing mechanisms are established to inhibit the influence that the expectancy representation has on responding. This model is by no means sophisticated yet it makes a small contribution to explorative ideas concerning the interactions of different elements of a timing system in the framework of SET.

The influence of motivational or emotional processes that might be linked to response rate and sustained attention are not incorporated in our models or the standard model. The only model that incorporates these factors is the spectral timing model (Grossberg and Merril, 1992, 1998). There has been a sustained interest concerning the role of attention over the years, particularly the effect of information load in temporal perception (See Zakay and Block 1998). One of the major goals of developing a connectionist model in the framework of SET is to incorporate both dynamic atttentional and motivational components in to the model.

Neither the standard model or our model involve recurrent interactions between different elements of the network. Only the interval based models that have a much more neuroscientific flavour, involve such interactions. One may be able to envisage how a highly interactive neural network could influence response precision and response rate. For example the relativity of response precision may not be due to some learnt threshold as proposed in this model but may result from time dependent changes in output values produced by some mechanism that influences time estimation.

Last of all it is worthy noting the flexibility of the accumulator model as a research tool. It is not designed specifically for the peak procedure like the standard model. As the network provides a single summed value it can be used in conjunction with all kinds of decision rules for further experimentation.

10.2 Time dependent spatial patterns

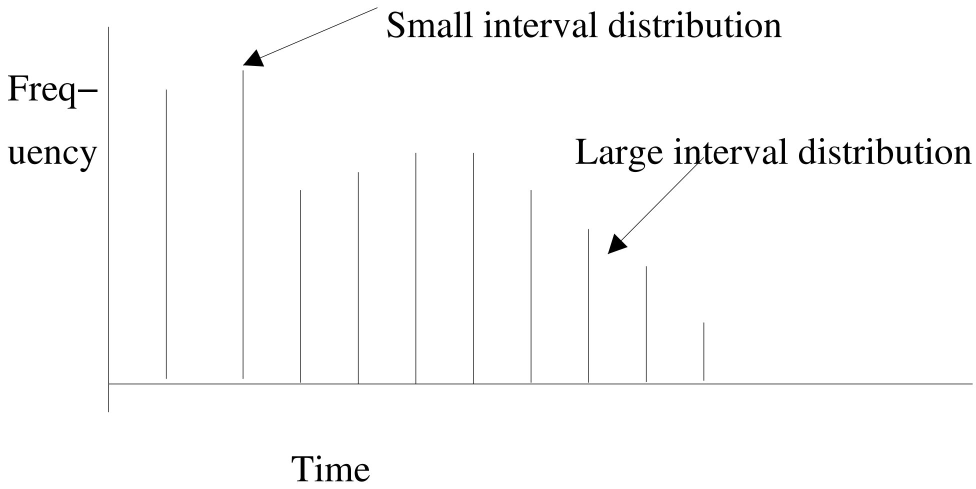
In section 6 the implicated role of the frontal lobes and basal ganglionic system in timing were briefly discussed. As mentioned the basal ganglionic system comprises of loops that feedback to the prefrontal cortex. Such recurrent feedback could be involved in generating time dependent changes in a population of neurons that code for time. This might be an interesting research avenue if not by way of neurobiological consideration then at least in very abstract networks.

Also discussed in section 6 is the problem of quantification with spatial representations of time. It may be the case that time−left performance does not rely on quantifiable processes but similarity comparisons of spatial patterns. As previously mentioned if one assumes a system generates time dependent spatial patterns in a way that permits the difference between any two spatial patterns to reflect the temporal differences which they represent, then tasks like time−left could be explained through association rather than quantification. Note however that the association of features may become none existent between very distant interval sizes in such a scheme. This could be investigated.

10.3 The relativity of differences

Apart from the speculations discussed in section 8, an alternative method could be explored for computing relative differences? In this scheme the accumulator generates time dependent spatial patterns that may or may not have a quantifiable structure. Learning in the system involves weights and/or threshold modifications that allow units to recognise the spatial pattern of the accumulator at some interval t. If the accumulator has scalar variance the memory system would be presented with a variety of spatial patterns at reinforcement time. Some of these patterns will occur with high frequency while others will not, reflecting a Gaussian distribution of such patterns. The heterogeneity of spatial patterns exposed to the learning system at reinforcement time will increase with the size of the interval being timed. In training a network to recognise patterns one would assume that the patterns that occur with the highest frequency influence learning the most. The learning process must also consider the frequencies of different spatial patterns. To understand this, see illustration 10.1. There are two accumulator distributions. The vertical lines separate the spatial patterns in to categories for ease of understanding. The four centre categories of the larger interval distribution have little difference in frequency, whereas frequency differences as a function of time for the small interval distribution is much more marked. A learning rule in this scheme must be sensitive to these differences to produce relative response estimations.

Illustration 10.1 Accumulator distributions of spatial patterns



The ability of the memory system to discriminate the most frequent patterns (which represents the reinforcement time) will be attenuated the more heterogeneous the distribution of spatial patterns at reinforcement time. Therefore one might envisage a network that generalises more when the spatial patterns of a distribution are more heterogeneous. As this heterogeneity increases with the reinforcement time one would predict greater generalisation as a monotonic function of the reinforcement time. In addition, the recognition system will generalise patterns above and below the mean pattern at reinforcement time. Hence the idea of absolute difference computations can be done away with. In fact, in this scheme, the idea of autonomous decision processes would no longer exist.

This is only speculation yet is may be worthy of preliminary experimentation with discrimination architectures like perceptrons. This might also be interesting in relation to bisection performance where a perceptron is trained to discriminate between S and L accumulator values. Note that while perceptrons may be worthy of initial experimentation, as discussed, one would need to construct a dynamic connectionist process to provide a more plausible account of timing. In short one needs to develop a dynamic neural network that both classifies data it experiences at reinforcement time and generalises as a function of a measure of the dissimilarity of these experiences.

One problem with such a scheme is that the correct degree of temporal generalisation would depend on the amount of training trials experienced. Animals can perform transitions from one time to another in a very small number of trials. The author is unclear whether single trial performance immediately following a novel transition maintains the relative response precision. This could be examined.

A group of architectures, which involve 'self organised' learning styles, could also be experimented with in a different way. These kinds of architectures classify patterns that share statistical features. One could experiment with parameters that affect the broadness of pattern classification. In a network these parameters could be manipulated to change as a function of elapsed time. Such a scheme could reflect a hard−wired rather than a learnt relative generalisation process.

10.4 Summary

In this thesis an explorative approach has been adopted to connectionist modelling of animal timing. The models proposed have been very abstract but this is justified by the lack of neurobiological understanding of timing processes. The simple networks employed generate data that approximates data reported in the literature. The model has attempted to experiment with or discussed three central features of SET:

1) Scalar variance in the accumulator rather than memory

2) The relativity of response calculations based on a generalisation parameter relative to the heterogeneity of accumulator distributions rather than explicit ratio calculations.

3) Geometrical mean performance as a result of discrimination between overlapping accumulator distributions rather than asymmetrical ratio

calculations.

There are also a number of limitations of the neural networks that have been reported. These include:

1) Level of abstraction employed.

2) Does not account for hedonic or motivation factors that may be implicated with attentional processes and response rate.

3) A discrepancy between the kind of memory process used in this model and the dynamic attractor like memory process implicated by experimental data.

4) Inability to account for data observed when transitions between different reinforcement times are made.

5) A discrepancy between BDH and experimental data suggesting explicit representations are employed.

Connectionist modelling of temporal cognition in tasks accounted for by SET is clearly a neglected research area. This is really quite surprising considering the highly quantitative nature of the research field. This research area is certainly ideal for connectionist modelling for a number of reasons. Firstly there is the diversity of reliable quantitative patterns associated with different experimental manipulations. Secondly, there is the elusiveness of some of the patterns observed (e.g. geometrical mean bisection). Thirdly, it is highly likely that these patterns emerge form dynamic interactions between different cognitive and affective processes. Last of all, SET would appear to require a more contemporary reconceptualisation. Although the mathematical rules of SET explain data very well, the sequential information processing view of SET has little intuitive plausibility in the context of current psychological thinking and the prevailing dynamical view of natural cognition (See Port and Van Gelder, 1995). The standard model is considered a connectionist model, yet it's none dynamical nature (as with the model reported here) places it very close to the original information processing theory. There is wide scope for research development in this field!

References

Buonomano, D. V. & Merzenich, M. M. (1995) *Temporal information transformed into a spatial code by a neural network with realistic properties.* Science, **267**:1028–130.

Buonomano, D.V. & Mauck, M.D. (1994).*Neural network model of the cerebellum: Temporal discrimination and the timing of motor responses.* Neural Computation, **6**: 38–65.

Church, R, M.(1997). *Timing and temporal search.* In C.M Bradshaw & E. Szabadi (Eds.), Time and Behaviour: Psychological and Neurobehavioral Analyses (pp459–516). Elsevier.

Church, R. M. & Broadbent, H.A. (1990). *Alternative representations of time, number, and rate*. Cognition, **37**:55–81.

Church, R.M. & Deluty, M.Z. (1977). *Bisection of temporal intervals*. Journal of Experimental Psychology: Animal Behaviour Processes, **3**: 216–228.

Church, R.M. & Gibbon, J. (1982). *Temporal generalisation*. Journal of Experimental Psychology: Animal Behaviour processes, **8:**165–186.

Churchland, P.S. & Sejnowski, T.J. (1992). *The computational brain*. The MIT Press.

Gibbon, J. & Church, R. M.(1981). *Time Left: linear versus logarithmic subjective* *time*. Journal of Experimental Psychology: Animal Behaviour Processes, **7**:87–106.

Gibbon, J. & Church, R. M.(1990). *Representation of time.* Cognition, **37**: 23–54.

Gibbon, J. Malapani, C., Dale, C.L. & Gallistel, C. R. (1997). *Toward a neurobiology of temporal cognition: Advances and challenges,* Current Opinion in Neurobiology, **7**:170–184.

Gibbon, J., Church. R. M. & Meck, W.H.(1984). *Scalar timing in memory.* Timing and Time perception. Annals of the New York Academy of Sciences, **23**:52–77

Grossberg, S and Merrill, J.W.L (1992). *A neural network model of adaptively timed reinforcement learning and hippocampal dynamics,* Brain Research, **1**: 3–38.

Grossberg, S. & Merrill, J.W.L. (1998). *The hippocampus and cerebellum in adaptively timed learning, recognition and movement.* Journal of Cognitive Neuroscience, **8**: 257–277.

Harrington, D.L., Haaland, K.Y. & Hermanowicz, H. (1998). *Temporal Processing in the Basal Ganglia*, Neuropsychology, **12** :3–12.

Hinton, S.C. & Meck, W.H. (1997). *How time flies: Functional and neural mechanisms of interval timing*. In C.M Bradshaw & E. Szabadi (Eds.), Time and Behaviour: Psychological and Neurobehavioral Analyses (pp 459–516). Elsevier.

Houk, C.J. (1995*). Information processing in modular circuits linking the basal ganglia and cerebral cortex.* In C Houk, J L Davis. and D G Beiser (Eds.), Models of information processing in the basal ganglia (pp3–10).The MIT Press.

Ivry, R. B. (1996).*The representation of temporal information in perception and motor control.* Current Opinion in Neurobiology, **6**:851–857.

Ivry, R.B., Keele, S.W & Diener, H.C (1988). *Dissociation of the lateral and medial cerebellum in movement timing and movement execution.* Experimental Brain Research, **73**: 167–180.

Beale,R. & Jackson,T. (1990). *Neural computing*. J W Arrowsmith Ltd.

Jeffress, L.A.(1948). *A place theory of sound location*. J. Comp. Physio. Psychol. **41**:35–29. Cited in Miall (1996).

Kesner, R.P. (1998). *Neural mediation of memory for time: Role of the hippocampus* *and medial prefrontal cortex.* Psychonomic Bulletin and Review. **5**: 585−596.

Lejueune, H., Ferrara, A., Simons, F., & Wearden, J.H. (1997*). Adjusting to changes in the time of reinforcement: Peak−interval transitions in rats*. Journal of Experimental Psychology: Animal Behaviour Processes, **23**:211−231.

Macar, F. & Casini, L. (1998). *Brain correlates of time Processing*. In V. De Keyser, G. d'Ydewalle, & A. Vandierendonck (Eds.), Time and the Dynamic Control of Behaviour. (pp71−82) Hogrefe & Huber Publishers.

Malapani, C., Rakitin, B., Levy, R., Meck, W.H., Deweer, B., Dubios, B. & Gibbon, J. (1998). *Coupled temporal memories in Parkinson's disease: A dopamine−related dysfunction*. Journal of Cognitive Neuroscience **10**:316−331.

Meck, W.H (1996). *Neuropharmacology of Timing and Time Perception*. Cognitive Brain Research, **3**: 227−242.

Meck, W.H., Komeily−Zadeh, F.N., & Church, R.M.(1984).*Two−step acquisition: Modification of an internal clock criterion.* Journal of Experimental Psychology: Animal Behaviour Processes, **10**: 297−306.

Miall, R.C. (1989). *The storage of intervals using oscillator neurons*. Neural Computation, **1**: 359−371.

Miall, R.C. (1992). *Oscillators, predictions and time*. In F. Marcar, V. Pouthas & W.J. Freidman, (Eds.) Time, Action and Cognition: Towards Bridging the Gap (pp215−227). Kluwer Academic Publishers.

Miall, R.C. (1993). *Neural networks and the representation of time*. Psychol. Belg. **33**: 255−269.

Miall, R.C. (1996). *Models of Neural Timing.* In M .A. Pastor & J. Artieda (Eds.) Time, Internal clocks and Movement (pp 69−94). Elsevier.

Moore, J,W. (1992) *A mechanism for timing conditioned responses.* In F. Marcar, V. Pouthas, & W.J. Freidman, (Eds.) Time, Action and Cognition: Towards Bridging the Gap (pp229−238). Kluwer Academic Publishers.

Nichelli, P., Always, D. & Grafman, J. (1996). *Perceptual timing in cerebellar degeneration*. Neuropsychologia, **34**:863−871.

Niki, H. & Watanbe, M. (1979). *Prefrontal and cingulate unit Activity during timing behaviour in the monkey*. Brain Research, **171**:213–224.

O'Boyle, D.J Freeman, J.S. & Cody, F.W.J. (1996). *The accuracy and precision of timing of self–paced, repetitive movements in subjects with Parkinson's disease*. Brain, **119**:51–70

O'Boyle, D.J. (1997). *On the human Neuropsychology of timing of simple repetitive movements.* In C.M Bradshaw & E. Szabadi (Eds.), Time & Behaviour: Psychological and Neurobehavioral Analyses (pp 459–516). Elsevier.

Olton, D.S., Wenk,G.L., Church, R.M. & Meck, W.H.(1988). *Attention and the frontal cortex as examined by simultaneous temporal processing.* Neuropsychologia, **26**:307–318.

Ornstein, R.F. (1969). *On the experience of time*. Harmondsworth, England: Penguin.

Parkin, A.J. (1996). *Explorations in cognitive neuropsychology*. Blackwell Publishers Ltd.

Platt, J.R & Davis, E.R.(1983). *Bisection of temporal intervals by pigeons.* Journal of Experimental Psychology: Animal Behaviour Processes. **9**:160–170.

Quinlan, P.(1991).*Connectionism and psychology: A psychological perspective on new connectionist research*. Harvester Wheatsheaf.

Roberts, R.(1981). *Isolation of an Internal Clock*. Journal of Experimental Psychology: Animal Behaviour Processes, **17**:342–368

Rodriguez–Girones, M. & Kacelnik, A. *Response latencies in temporal bisection: implications for timing models*. In V. De Keyser, G. d'Ydewalle, & A. Vandierendonck (Eds.), Time and the Dynamic Control of Behaviour. (pp51–70) Hogrefe & Huber Publishers.

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

Schugens .M,M., Daum, I., Ackerman, H., Lutzenberger, W. & Birbaumer, N. (1994). *Processing of temporal intervals in patients with cerebellar degeneration*. Journal of Psychophysiology, **8**:366–367.

Siegel, S.F & Church, R.M. (1984). *The decision rule in temporal bisection*. Annals of the New York Academy of Sciences. **423**:643−645.

Siegel, S.F. (1986). *A test of the similarity rule model of temporal bisection*. Learning and Motivation, **17**:59−75.

Treisman, M. (1963). *Temporal discrimination and the indifference interval: Implications for a model of the internal clock*. Psychological Monographs, **77**: 1−31.

Wearden, J. H. & Doherty, M.F. (1995). *Exploring and Developing a connectionist Model of Animal Timing: Peak Procedure and Fixed−Interval Simulations.* Journal of Experimental Psychology: Animal Behaviour Processes, **21**:99−115

Wearden, J. H. (1999*). "Beyond the Fields we Know..": exploring and developing scalar timing theory,* Behavioural Processes (In press).

Wearden, J.H. (1994). *Prescriptions for models of Biopsychological time*. In M. Oaksford & G. Brown (Eds.), Neurodynamics and Psychology (pp. 215−236). San Diego, CA: Academic Press.

Wearden, J.H., & Ferrara, A. (1996). Stimulus range effects in temporal bisection by humans. Quarterly Journal of Experimental Psychology, **48:**24−44.

Zakay, D. & Block, R. (1998). *New perspectives on prospective time estimation*. In V. De Keyser, G. d'Ydewalle, and A. Vandierendonck (Eds.), Time and the Dynamic Control of Behaviour. (pp129−142) Hogrefe & Huber Publishers