Introduction [to Part II: Tapping and Synchronization].

John A. Michon

In the 1960s a modest revolution took place in the study of visual perception that had a vague but meaningful connection with the wave-particle duality of light in physics. It became clear that there would be an advantage in looking at phenomena such as visual contrast, acuity, color, and shape perception, not only from the conventional object-oriented point of view, but from a spatial frequency perspective as well. A new line of research emerged: the traditional measurement of visual acuity by means of Snellen charts and Landolt rings, for instance, was supplemented by having people stare at displays of sinusoidally light/dark modulated spatial ‘gratings’, making acuity amenable to dynamic systems analysis. This approach has been reviewed in detail by Olzak and Thomas (1986).

At roughly the same time a reverse development took place in auditory perception, where frequency had always been the natural mode of representation. Initially stimulated by J. J. Gibson’s (1966) ideas about ecological optics and ‘affordance’, the concept of ‘acoustic object’ penetrated into auditory psychophysics. A major role in this development was played (by Bregman, 1981, 1990). As a result of Bregman’s work on auditory scene analysis we have gained a deep understanding of the relation between a musical object, say a bassoon, and the stream of auditory information that specifies the essence of what makes a sound pattern emerge as the ‘scene’ of a bassoon-being-played-on, standing out perhaps even in the middle of the dense pattern of sound waves that is produced by a symphony orchestra.

This research illustrates that the simple relation between period and frequency, $T = 1/f$, lies at the root of a deep phenomenological distinction characterizing the dual nature of temporal experience, the distinction, that
A cognitive framework

In the past most studies on the perception and production of brief temporal intervals followed the tradition of the psychophysical school of time psychology (Michon and Jackson, 1985). Only in the course of the last decade or two have we seen a gradual shift away from the strict paradigms of temporal psychophysics towards a more explicit cognitive approach. Several variations on popular mainstream cognitive themes can be distinguished, including dynamic attending (Jones and Boltz, 1989; Large and Jones, 1999), perceptual dynamics (Freyd, 1987; Freyd, Kelly, and DeKay, 1990), and code theory (Leeuwenberg and Buffart, 1978; Van der Helm, 1988). But even the work that stays close to the original tradition is undergoing such a change of perspective, suggesting that the dual nature of temporal experience has taken hold at last in the domain of cognitive psychology.

The received view of human cognitive activity has emphasized a distinction between automatic and consciously controlled information processing (Schneider and Shiffrin, 1977) although more recently many have come to prefer the terms implicit and explicit instead (Schacter, 1987, 1996). The temporal information that is being processed in either mode may be contained either in the real time perceptual input or in the simulated time input that resides in memory. This suggests $2 \times 2$ modes of information processing (see Figure 1). The differences in outlook provided by each of these processing modes are suggested by comparing such characteristics as the basic conceptual ‘chunks’ involved, the temporal correlates, formal processing characteristics and conceptual analogies that are used to describe and analyse a temporal input.

When these distinctions were first made in the mid-seventies, the possibilities of incorporating them into ‘deep’ theory were quite limited. Most models were formulated in a qualitative fashion. That is, research aimed primarily for results that would be at least qualitatively consistent with assumptions about the mechanisms underlying a particular phenomenon, be it streaming, anticipation, or synchronization. More recently, however, the rapid progress in computing power has led to an unprecedented growth of the experimental and formal sophistication of studies in music, time, and rhythm, allowing a much more quantitative approach that does better justice to the ‘dual nature’ of time experience.
### Tapping and Synchronization

<table>
<thead>
<tr>
<th>Information Processing</th>
<th>Automatic</th>
<th>Controlled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Inputs</td>
<td>Perception of dynamic invariants</td>
<td>Perception of objects and events</td>
</tr>
<tr>
<td>Encoded Inputs</td>
<td>Procedural or implicit knowledge</td>
<td>Declarative or explicit knowledge</td>
</tr>
<tr>
<td>Conceptual 'chunks'</td>
<td>Affordances</td>
<td>Episodes</td>
</tr>
<tr>
<td>Temporal Correlates</td>
<td>Frequency phase cycle</td>
<td>Duration order date</td>
</tr>
<tr>
<td>Processing Characteristics</td>
<td>Parallel direct distributed 'data-driven'</td>
<td>Sequential computed 'rule-driven'</td>
</tr>
<tr>
<td>Conceptual Analogies</td>
<td>'Implicate' implicite holistic ... field resonance wave</td>
<td>'Explicate' explicite focal ... problem space computation element</td>
</tr>
</tbody>
</table>
There remains a good deal of conceptual continuity. Despite all methodological progress the theoretical issues that motivated much of the research in the perception and production of rhythm in the 1960s seem to still be around and appear far from solved. This section on Tapping and Synchronization is illustrative in this respect, not only because of the concrete results the authors present, but also because their presentations reflect the changing ways in which researchers tackle the questions how people perceive and produce series of isochronic or patterned intervals and why they deviate from strict regularity in doing so. These questions had already drawn my attention early in the sixties, at a time when finger tapping and time perception studies enjoyed a modest popularity which gained momentum with the publication of Fraisse’s landmark monograph Psychologie du Temps (Fraisse, 1957), but perhaps to an even greater extent as a result of the studies of Frankenueuser (1959) and Treisman (1963). My own contributions at the time were the introduction of timing as a means of measuring the ‘mental load’ of various perceptual-motor tasks (Michon, 1966) and the description of timing behavior in terms of linear dynamics (Michon, 1967).

In those days the tools of non-linear dynamics were still far from being applicable to the study of complex skill, partly because the computational means required for numerical solutions of non-linear equations that we now have at our disposal were nowhere in sight. A serious further limitation of linear dynamics is that any factor one cannot force to linearity by appropriate experimental constraints (e.g., ‘small signal linearization’) must be subsumed under the residual terms of one’s equations. Non-linear dynamics allows a treatment of various factors that influence performance but that are otherwise difficult to control experimentally.

**Limits of stationarity**

This development makes it possible, for instance, to begin discussing the cognitive origins of drift, which I see as the principal merit of Madison’s contribution (even though the author appears to prefer a more formal methodological perspective). Madison takes a close look at the fact that in producing a sequence of intervals subjects will, almost without exception, drift away from the intended rate. He presents evidence indicating that this drift may be described in terms of ‘pink’ or $1/f$ noise, which would seem to be compatible with an explanation of these results in terms of biological mechanisms. As some of the author’s results seem to be better in line with $1/f^2$ than with $1/f$ (see his Figure 3 on p. 107), an explanation that is
consistent with the latter type of noise would perhaps offer an interesting and behaviorally plausible alternative. This so-called ‘brown’ $1/f^2$ noise describes a mechanism—of which Brownian motion constitutes the prototypical example—of the kind one would expect to operate in this case. Subjects may well choose a standard based on some of their most recent performance and then display random variations relative to that standard. Characteristically the fluctuations of this standard will be slow relative to the variations of the actual series.

Whilst empirical evidence to support this position is limited, the thought as such has occasionally surfaced during the last 25 years. Thus, already in 1972, Vroon reported a series of temporal tracking experiments focusing on the drift phenomenon (which, at the time, he called the ‘lengthening effect’). On the basis of his findings he was able to rule out several explanations. Among other things he established that drift is independent from the intensity of the auditory stimulation (introduced as a way of manipulating activation level) and from ‘resetting’ instructions. These and similar findings carried him to the conclusion that:

“...the central factor is the estimation process itself. The data lead to the conclusion that the last trial has to be considered as the only basis of lengthening. The last estimation is continuously experienced as too long or too short, and the resulting corrections lead to the described compensatory actions” (Vroon, 1972, p. 233)

Vroon’s results do indeed suggest a mechanism that is consistent with the local push-and-pull nature of ‘brown’ noise and that depends very much on local information affecting decision making, memory strategies and sequential response biases (Vroon, 1972, p. 234).

Recently Panissal (1998), looking for a “supra-modal mechanism underlying the processing of interval information,” reported results that may be considered consistent with this view. Although Panissal made an effort to attribute her results to a single mechanism, it is a likely possibility that the severe experimental constraints she imposed on her subjects had forced them to act as if their behavior was driven by one timing mechanism only. In addition, only highly experienced subjects acted consistently with her single-mechanism hypothesis. In contrast, the results of naive subjects were better in line with a two- or three-mechanism hypothesis. This raises the question who are, in fact, displaying normal temporal abilities of the human species: is it the naive subjects or the highly trained temporal experts? Can it be that in normal everyday life two or three timing mechanisms are active that are eventually superseded by a single mechanism, as a result of a long and involved period of interval training and the constraints imposed by the experimenter? And, still in this context, when and how does interval-related information give way to frequency and phase-related information?
Limits of synchronization

In their contribution Wohlschläger and Koch continue this debate, more specifically in the context of what they call the synchronization error, a specific aspect of tuning performance. When a person produces a series of regularly spaced stimulus taps in synchrony with a pattern of clicks, subjects tend to tap 20-60 milliseconds before the click with which they are synchronizing. One may wonder whether this really should be considered a synchronization error: professional musicians make elaborate use of it in a highly controlled and systematic fashion (e.g., Rasch, 1981). As before, several factors may influence this result: the pacing signal, tempo, effector organ, feedback, and practice. Wohlschläger and Koch review various attempts at explaining the nature of the negative synchronization error: the P-center hypothesis, the afferent-efferent conduction time difference hypothesis proposed by Paillard and Fraisse, and the sensory accumulation model suggested more recently by Gehrke, Aschersleben, and Prinz (1998). The latter resembles the explanation offered by Paillard-Fraisse, but it rests on an additional assumption regarding the integration of sensory information. As a result the sensory accumulation model should be able, among other things, to explain that intense taps lead to a reduced negative synchronization error, a finding that seems to be at an angle with Vroon’s (1972) finding, mentioned above, that stimulus intensity does not affect timing performance. On the basis of their results, however, Wohlschläger and Koch come to the conclusion that neither of these three explanations by itself is sufficient to explain the negative synchronization error. They summarize their position by stating that “it might be that the synchronization error is an effect that occurs under artificial laboratory conditions, in other words an artifact.”

Limits of phase correction

Repp raises the T vs. f question in the context of phase error correction. He concludes that “phase errors are continuously compensated for, even if they fall below the temporal order threshold and regardless of whether they are due to a stimulus perturbation or to response variability alone.” The temporal order threshold is thought to impose limits only on conscious perception, not on the processing of temporal information for motor control. This claim is substantiated by assuming that there may be a preconscious stage of temporal perception that is directly coupled by entrainment to motor control mechanisms. This is an assumption that has come to play a central role in Large and Jones’ recent analysis of dynamic attending (Large and Jones, 1999).

But thresholds being what they are—ranges of uncertain outcomes is what they are—how is the temporal order threshold to be valued? Since
the days of Hirsh and Sherrick (1961) it has been fixed at approximately 25 milliseconds. Yet, if we take more of the cognitive ‘surroundings’ of our experiment into account again, then the question arises, to what extent the experimental conditions play a role in Repp’s results. Warren and his associates (Warren, Obusek, Farmer, and Warren, 1969; Warren and Obusek, 1972), for instance, demonstrated that linguistically and musically meaningful sounds show vastly lower temporal order thresholds than meaningless sounds such as beeps, screeches and hisses. In the latter case the temporal order threshold even turns out to be of the order of 150 milliseconds.

Importantly for our present discussion, Repp argues that frequency and phase perception are essentially different from duration perception. And as he points out, an unanswered question thus far is whether there is a smooth transition from one mode of perception to the other. Such a transition constitutes the conventional assumption. Fraisse, for instance, suggested a perceptual shift from rhythm perception for rates above 2 Hz as opposed to interval perception for periods of 1 second and more (Fraisse, 1957). The alternative would be that mechanisms for period perception and frequency/phase perception are complementary. In that case the choice of perceptual mode would depend on the nature of the temporal task. As Repp argues, this requires experiments that “question both modes for the same stimuli.”

**Limits to temporal ratio**

Franěk, Mates and Nátová focus on an ancient question that played an important role already in Saint Augustine’s incisive analysis of experiential time: Why is it that rhythmic groups regress to simple combinations of ‘long’ and ‘short’ intervals and how do we compare these?

So far as sense can make things plain I measure a long syllable by a short, and I feel by means of my senses that it has twice the length. But when two syllables sound one after the other—the first short, the next long—how shall I keep hold of the short one? How in my measurement shall I apply it to the long one, so as to find that the long one has twice its length? […] I cannot make this judgment except when both the syllables have gone into the past and are finished. Therefore, what I am measuring is not the syllables themselves (they no longer exist) but something in my memory which remains there fixed. (Augustine, c. 400/1972, bk XI; 27).

Phenomenally the 2:1 ratio turns out to be almost indestructible. Even groups that are widely off this mark, such as 5:1:1 for instance, can be perceived or (re)produced as a simple 2:1:1 group. Altogether it seems clear that stable representations of more complicated ratios depend on higher order perceptual and intrinsic pattern structure (e.g., Garner, 1962, 1974; Povel, 1981, 1985).
Franěk et al. come to grips with this problem by forcing their subjects to either double or halve their speed of finger tapping. They find systematic deviations from the exact rates that support earlier findings related to the indifference interval (a range of greatest precision around 500 ms) and the time-order error (overestimation of short intervals and underestimation of long intervals, relative to the indifference interval). The authors offer a range of possible explanations for these results, but also make it clear that further experiments, especially of hypothesis-falsifying kind, are needed.

The authors also generalize Saint Augustine’s question in an interesting way. They argue that a tempo stored in the memory of, say, a musician, is likely to influence a timing task but that experimental psychology has failed to study such influences. I concede that it is important to study the effects of stored (encoded) temporal relations on actual inputs, but I am less convinced than the authors that this is “an entirely unknown area” of psychological investigation. It should be rewarding to look for inspiration in areas like motor skills, natural computation and non-linear dynamic modeling.

**Robot in search of a time sense**

Of the various contributions in the section on Tapping and Synchronization, the contribution by Eck, Gasser and Port is most closely related to the approach that takes its point of departure in non-linear dynamics. Their paper raises a number of fundamental issues relating to modeling in general, issues that stood out in pretty much the same way in the ancient linear-dynamics approach I mentioned earlier.

The authors claim that the process of learning to be rhythmical can be best studied by building a robotic model. In one important sense I agree as

“...one should eventually be able to provide the design specifications for an intelligent system that, by dealing with a dynamic environment, could succeed in timing its mind and also, by explicitly manipulating its temporal experience, might be said to be minding its time” (Michon, 1989, p. 19)

From the cognitive perspective such a robot should be constructed on the basis of assumptions about the functions that can generate its (goal-directed) behavior, rather than representing it, in bottom-up fashion, as a Turing-style match between a real person’s output and the robot’s output. This requires an a priori computational theory sensu Marr, (1982) or, similarly, a rational analysis sensu Anderson (1993; see also Oaksford and Chater, 1998).

The dynamic model Eck, Gasser and Port present in their contribution consists of a mass-spring robot arm that is driven by a small
network of ‘neural oscillators’ with fast coupling, leading to entrainment. They use a weak heuristic for model building, one that essentially consists of selecting properties one by one that mimic an aspect of the empirical phenomenon. This bottom-up strategy induces them to consider the perception and production of rhythm as the prototypical ‘model’ for handling temporal information. This is the classical engineering approach to ‘systems identification’ which, unfortunately, without further functional constraints offers no guarantee that the selected properties will work in combination and do all the things one is hoping for and — importantly — only these things.

Also the idea that beat induction would qualify as a benchmark domain for studying the timing interactions between brain, body, and the external world may be too optimistic. As I have argued many times before, I think that staying tuned to the dynamic events in the world around us is the ultimate foundation of our temporal existence which, as such has a very long and essential evolutionary history (e.g., Michon, 1985, p. 20). Falling out of tune with the world is simply and instantly fatal. It is with reference to this tuning function that a rational analysis must be carried out. Elements of such analysis are emerging in the literature; relevant examples, in addition to some chapters in the present volume, are Brown and Vousden (1998), Large and Jones (1999), Large and Kolen (1999) and Port, Cummins, and McAuley (1995), to name just a few that are directly relevant to the present discussion.

References


TAPPING AND SYNCHRONIZATION


