jisaki and Kawashima (1970) proposed a model explaining the decision process based on continuous auditory information. No reader who was unfamiliar with the literature on speech perception would suspect that this conclusion has been a commonplace of that literature for nearly 20 years, and was indeed well established before Massaro even entered the field.

Established before Massaro even entered the field.

In fact, the hypothesis of Studdert-Kennedy et al. (1970) that stop consonants (though not vowels) could be discriminated only if they were differently identified had already been undermined before it was published. Stevens (1968) and Sachs (1969) both attributed the poor within-category discrimination of stop consonants to the transient nature of their acoustic cues. Fujisaki and Kawashima (1970) proposed a model explaining the consonant-vowel difference as due to differential decay of auditory memory during a discrimination trial. Pisoni (1973) and his colleagues (Pisoni & Lazarus 1974; Pisoni & Tash 1974) supported Fujisaki’s “dual process” model in a series of elegant and decisive experiments to which Paradigm gives no credit. Studdert-Kennedy and Shankweiler (1970) also attributed the consonant-vowel difference in dichotic ear advantages to differential loss of auditory information during transcortical transfer.

Moreover, all these studies and many others have been repeatedly acknowledged in review articles (e.g., Repp 1984; Studdert-Kennedy 1975, 1976, 1980, 1981). For example: “In short, consonants and vowels are distinguished in the experiments we have been considering, not by their phonetic class or the processes of assignment to that class, but by their acoustic characteristics and by the duration of their auditory stores” (Studdert-Kennedy 1976, p. 263). Compare this statement with the following from Paradigm: “If the sensory information is lost very quickly, continuous information could participate in the perceptual process but might not be readily accessible for introspective reports” (p. 120, p. 14 of target article). Massaro apparently does not recognize the correspondence between these two statements. But then he does not acknowledge many of the papers cited in this and the preceding paragraph, nor the many other relevant papers cited in those papers. Perhaps he has not read them. If so, one can only observe that those who do not study the literature are condemned to repeat it.

Finally, let me note that Paradigm grossly underrates the importance of categorical perception in support of the hypothesis that speech perception engages specialized cerebral processes. The principal supports for this hypothesis are: (1) the failure, despite attempts over many years, to devise an arbitrary acoustic substitute for speech in reading machines for the blind that would be any more efficient than Morse code, for which perceptual rates are scarcely one tenth those of speech (Liber & Massaro 1977); (2) the repeated demonstration, in both brain-damaged and normal subjects, of a double dissociation between left and right cerebral hemispheres in the perception of speech and nonspeech sounds (e.g., Blumstein 1981; Molles & Betz 1988; Tartter 1988). That such findings have no place in the grand design of a Paradigm for Psychological Inquiry invites the conclusion that this is a paradigm we would do well to lose.

Acknowledgment

Preparation of this comment was supported in part by NIMH Grant 20924 to Haskins Laboratories.

Note

• Author’s additional affiliations are University of Connecticut and Yale University.

Winning “20 Questions” with mathematical models

James T. Townsend

Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907

Electronic mail: townsend@brazil.psych.purdue.edu

Massaro’s book is a milestone in perceptual/cognitive psychology. First and foremost, it shows that: (1) The information processing approach is not only not dead, but still has much to offer behavioral science. (2) It is possible to test psychologically interesting opposed or dichotomous alternatives (i.e., it is possible to successfully play the game of “twenty questions”, see e.g., Newell 1973) using creative, theoretically motivated experimental designs. (3) Mathematical modeling offers great power in this quest: (4) The discussion of such issues as modularity, though enriched by philosophical as well as physiological disputation (e.g., Fodor 1983), requires behavioral intervention for successful resolution.

All this having been said, and accepting that Massaro has amassed impressive evidence in favor of his claims, a number of the very important issues broached by Massaro are far from settled. For one thing, psychology’s experience is inevitably that as investigators dig deeper into phenomena in a search for lawfulness “things are always more complex than they seemed.” This has led again and again to the wholesale abandonment of topics (not just dichotomous issues either!) and research areas and has contributed to psychology’s reputation of flightiness and trendiness. This problem is not one we are likely to remedy in the near future, however.

Other problems are more immediate. One is that it is extraordinarily difficult to obtain answers from nature in a model-free way (as Massaro notes on p. 152). That is, for many research questions, it would be optimal if we were able to test all models which capture the critical psychological notions in which we are interested. This is rarely feasible, for many reasons. Even the testing of two dichotomous models of the same phenomena can be pragmatically impossible within the standard paradigms. For example, Luce’s choice theory (1959) predicts receiver operating curves that are, for all practical purposes, indistinguishable from the traditional theory of signal detectability (Swets et al. 1961). Parallel-serial identifiability difficulties are another example (Townsend 1971, 1972), although with effort, solutions can be discovered that permit experimental testability (e.g., Townsend 1976a, 1976b; Ross & Anderson 1981; Thomas 1969, Townsend & Ashby 1983). The main point here is that we are in most cases forced to start with a specific model or theory (that is, with assumptions and resulting predictions embedded in specific mathematical formulae) to test our hypotheses. It often takes years to work up the fully model-free tests and these may differ depending on the empirical facts learned in the meantime. (Note that putting models in verbal form does not solve this difficulty; it merely hides it.

This is perhaps the weakest link in Massaro’s armada of evidence. He is presumably interested in testing opposing concepts such as “categorical versus continuous” and “independent versus dependent” in some supramodel sense; and this, of
course, has not been done. Yet to the extent that his model is supported by many sets of data, both that model and its formalization of the various concepts (e.g., continuity) receive support. (Aside: I can’t help wondering why Massaro did not use a statistical technique such as the chi-square procedure which allows statistical tests to be performed along with the fits – as opposed to least squares, which don’t permit this in the present circumstances.)

Nevertheless, certain issues such as independence are particularly vulnerable to this criticism. A recent general theory of recognition (Ashby & Townsend 1986), which can be viewed as a generalization of signal detectability theory (Green & Swets 1966), portrays recognition quite differently from the way Massaro’s does. It also contains apparatus for response bias, an omission, though a remediable one, in Massaro’s theory. More generally, the way independence manifests itself in data could take many forms, depending on the particular model.

A final comment concerning the necessity of distinguishing the various types of independence (Ashby & Townsend 1986): In Chapter 6, for example, Massaro uses the notion of perceptual independence largely as Ashby and Townsend use it (e.g., in Massaro’s horse-race model of reaction time), whereas he later invokes a concept that we would refer to as “perceptual separability” (e.g., pp. 166–69). The various types of independence may be related in a particular model of perception, but they are logically distinct concepts.

**The use of mathematical models in perceptual theory**

Richard M. Warren

Department of Psychology, University of Wisconsin-Milwaukee, Milwaukee, WI 53201

Massaro states that “My goal in this essay is to integrate the information processing approach with the study of the world of information” (Paradigm, p. 10). In implementing this goal, he uses bimodal speech perception to develop a general model, or paradigm, for research in psychology. Hence, if this model is successful in dealing with the perception of speech, it would have a significance extending well beyond this particular topic. However, for the reasons given below, I believe that the ambitious goal of developing a generally useful paradigm has not been achieved.

Central to Massaro’s treatment of the role of audition and vision in speech perception is a series of seven alternative hypotheses which he calls “binary oppositions” (précis, Fig. 2; Paradigm, p. 24). These two-valued alternatives are used to construct “trees of wisdom” (Paradigm, Fig. 3). Choosing the appropriate alternative leads to another binary choice along the tree trunk (the inappropriate alternative is a dead branch, leading nowhere). Examination of the tree shows that the auditory and visual information concerning the speaker’s utterance can be either integrated or not integrated – then, if integrated, it can be categorical or continuous – then, if continuous, . . . and so on. The use of a succession of binary oppositions produces a fragile model, since each stage assumes the validity of all prior stages. Also, Massaro’s choice of the correct alternative is, in some cases, controversial. For example, his conclusion that speech perception (whether auditory, visual, or bimodal) is continuous rather than categorical (p. 150) conflicts with the conclusions of many, if not most, of the people who have done research in this area (see Harward 1987).

While Massaro uses binary oppositions for his “trees of wisdom,” he changes to truth values which can vary continuously from zero to one when constructing his Fuzzy Logical Model of Perception (FLMP). Massaro emphasizes that “fuzzy truth” is different from probability. He gives as an example: “If we say that a whale is a fish to degree .2, that does not mean that there is a .2 probability that a particular whale is a fish. Rather it is true that a whale is a fish to a degree .2” (Paradigm, p. 180). The FLMP uses quantified descriptors which function as empirically determined constants. The number of such constants is quite large, varying from 11 (Paradigm, p. 168) to 48 (pp. 128–29).

It appears to me that the FLMP violates Ockam’s razor by multiplying entities unnecessarily. It also seems to violate an earlier old proverb which might be called “Aristotle’s constraint” (expressed in his Nicomachean Ethics, book L, chapter 3): Do not quantify inappropriately. By using many empirically determined constants, each of which corresponds to a “parameter” in his FLMP, Massaro found it possible to obtain good fits to empirical data. With up to 48 constants to adjust, Massaro turned to the computer program STEPPIT, which iteratively adjusted the value of each parameter so as to minimize the squared deviation of the observed and “predicted” points (I’ll return to this use of “predicted” shortly). Massaro stated that, “The outcome of the program STEPPIT is a set of parameter values which, when put into the model, come closest to predicting the observed results. Thus, STEPPIT maximizes the accuracy of the description of each model” (Paradigm, pp. 19–20).

It has been stated, possibly with some exaggeration, that four constants suffice for plotting the figure of an elephant, and the introduction of a fifth can lift the tail. Consider what might be described with the numbers of constants used in the FLMP. Obtaining a fit between the output of the STEPPIT program and any experimentally obtained function would not seem surprising. Since this agreement is obtained after fixing the constants to match the empirical data, the term “predicted” does not seem appropriate. Yet the curve fitting results of the STEPPIT program are considered as a validation of predictions of the FLMP (see précis, Figs. 7 & 8).

Mathematical models have a seductive appeal. They permit experimental data to be described, not as a collection of individual values, but as measures that are linked by derivation from the same mathematical expression. However, if the agreement of the mathematical model with experiment is achieved only after the data are gathered and several constants are adjusted to fit, then this agreement would have little more significance for theory than the shaping of a flexible strip of rubber to fit an array of points on a graph.

**A comparison of speech perception and spatial localization**

Robert B. Welch

NASA-Ames Research Center, Mail stop 239-7, Moffett Field, CA 94035

With respect to both methodology and outcome, Massaro’s impressive research program on speech perception is highly reminiscent of studies of intersensory interaction in the perception of spatial location (Welch & Warren 1980). Both areas share the assumption that perception is based on the integration of multiple sources of sensory information whose relative weighting is most accurately assessed by placing them in conflict with one another. For example, momentarily viewing the hand through a prism causes it to feel as if it were located near (although not coincident with) its seen position (e.g., Hay et al. 1963), a phenomenon referred to, in general, as “intersensory bias.” From this observation it has been concluded that vision is more heavily weighted than proprioception in the everyday (i.e., sensorily congruent) perception of limb position. The present commentary is devoted to a discussion of the striking similarities between speech and spatial perception. This comparison is useful, first, because it strengthens Massaro’s contention that the perception of speech is not unique and, second, for its further illumination of this perceptual capacity and suggestions for future research.

Commentary/Massaro: Speech perception