Never published, withdraw from *JEP:General*, 8 June 92

On Perceptual Information and Fitting Models to Data: A Reply to Massaro and Cohen (1993)

James E. Cutting and Nicola Bruno

Cutting, Bruno, Brady, and Moore (1992) investigated the properties of Massaro's (1987) fuzzy-logical model of perception (FLMP) and other models. Among their results, they found that FLMP and an additive model fit the data of individuals' judgments of depth about equally well, that the additive model fit the group data better, and that FLMP fit random data better. Massaro and Cohen (1993) questioned the pertinence of the latter finding, questioned the fiting of the additive model (but not that of FLMP) to group data, and also compared our general approaches to perception. Here we show two new ways in which FLMP absorbs random error, raise a question about fits of FLMP to group data, and provide new support for Cutting et al's claims about the importance of fitting models to random data. We then make further comparisons between the approaches and explicate the concept of information specification which they queried.

Cutting, Bruno, Brady, and Moore (1992) investigated the perceived depth relations among three unmarked panels in a stimulus set defined by the orthogonal combinations of four sources of information--relative size, height in plane, occlusion, and motion parallax. They fit their data and those of Bruno and Cutting (1988) with several models including Massaro's fuzzy-logical model of perception (FLMP) and an additive model. Cutting et al replicated Massaro's (1988) result, finding little evidence that either of these two models was superior in capturing depth judgments of individuals.

Prompted by some anomalies in their results, Cutting et al then investigated the numerical properties of these models. They found that FLMP fit a wider variety of functions than did the additive model; indeed, it even fit random data better. After reanalyses and discussion they concluded (a) that there is still no reason to favor either FLMP or the additive model in judgments of depth, but (b) that to insure fair comparisons between models researchers should factor out the models' relative abilities to fit random data, and (c) that the relative complexity of FLMP, measured by its equation length, may account for its ability to fit many types of data and for its relative success in many empirical domains.

Massaro and Cohen (1993) presented little in direct response to the major conclusions of Cutting et al (1992), but they did present a number of important criticisms to other aspects of the Cutting et al paper. Among the conclusions of Massaro and Cohen were five concerning models and their results: (a) models are best tested against individuals, not groups, because of the problems inherent in averaging data across individuals and because of the issue of statistical regression which particularly favors the additive model, (b) the model-fitting results and analysis of variance results reported by Cutting et al were not anomalous or contradictory,[1] (c) FLMP is falsifiable,[2] (d) equation length is extraneous to any serious consideration of models,[3] and (e) the comparative fits of models to random data are irrelevant when comparing fits of those models to psychological data. In addition Massaro and Cohen concluded (f) that the paradigm and FLMP are closely related *to directed perception* (Cutting, 1986) but that certain aspects of the latter, such as the concept of *specification*, are ill-defined.

We divide our comments four ways. We concentrate on claims (a) and (e) above, discussing further the relationships between our group and individual data, and exploring further the ability of FLMP to absorb random error. Next we concentrate on two aspects of (f)--the similarities and differences in our approaches, and on explication of the concept of specification.

Group Data and Individual Data

Representativeness of Groups to Individuals

Massaro and Cohen (1993) were concerned with Cutting et al's (1992) statements about the relation between individual and group data. For example, Cutting et al had noted that the additive model fit the group data considerably better than did FLMP. Massaro and Cohen found this analysis unwise, and presented a familiar tutorial on how group data can be nonrepresentative of the individuals in the group. Indeed, we agree in principle with Massaro and Cohen; this is one of those methodological points about which psychologists all need to be reminded.

However, their tutorial is not relevant to the analyses of Cutting et al (1992). Cutting et al had moved to group analyses only after their previous analyses had shown that the individual data were, essentially, equally well fit by both models. The additive model was superior to FLMP in fitting about half of the individuals' data (23 of 44 subjects). Thus, the additive model and FLMP were both generally representative of individuals' data and, further, the comparison of fits of both models to the group data seems reasonable and appropriate.

After their tutorial Massaro and Cohen (1993) then presented a computational example of individuals, groups, and regression towards additivity. They generated ideal psychological data and demonstrated how FLMP can fit individual data better than the additive model, while at the same time the additive model fit the group data better than FLMP (see their Tables 1 and 2, and Figures 2 and 3). For purposes of pedagogy we wanted to replicate their analysis, but adapted to our situation.

New Simulations 1 and 2: Regression Does Not Always Favor the Additive Model

In Simulation 1 we selected median or representative functions from each of the four simulated individuals along Factor A in Figure 2 of Massaro and Cohen (1993). We then appropriated the general shape of each to fit our experimental situation with 0 through 4 sources of information. Our four functions are shown in Table 1. As in many of our previous simulations (Cutting et al, 1992), we fixed the value (with no variance) of perceived depth across the 4 stimuli with 1 source of information, the 6 stimuli with 2 sources of information, and the 4 with 3 sources of information; there was only one token each with 0 and with 4 sources of information. Thus, as before, the data of 16 stimuli were to be fit with the models.

We found, as did Massaro and Cohen (1993), that FLMP fit these hypotheticals individual data better than did the additive model, and that the additive model fit the hypothetical group data better than did FLMP. The only differences in our results and theirs were (a) that the advantage of FLMP over the additive model for the individual data, measured here in sum of least squares rather than RMSD, was not nearly so great as that shown by Massaro and Cohen for their situation, and (b) that the advantage of the additive model over FLMP for the group data was similarly smaller.

--Insert Table 1 about here--

Massaro and Cohen (1992, ms p.7) claimed that, due to statistical regression, the results of such simulations would be necessarily true in many cases, and that the reverse would never occur: "Averaging AMP [additive model of perception] results will always give ideal AMP results and will necessarily be more poorly fit by the FLMP than the AMP. A good fit of the FLMP to group results is meaningful whereas a good fit of the AMP is not." Unfortunately, these statements are not correct.

Shown in Table 1 for Simulation 2 are another set of hypothetical individual data. Notice, for the data of these individuals the additive model fits better than does FLMP, but that FLMP fits the group data better than does the additive model. This outcome is *exactly opposite* to Simulation 1. The existence of such a set of results contravenes Massaro and Cohen's claim that group analysis for FLMP is justifiable, but that one for the additive model is not.

Thus, we agree with Massaro and Cohen that analyses on group data must proceed with care. But we maintain that Cutting et al (1993) exercised proper caution, and we have demonstrated here that FLMP's fits to group data do not have the immunity to statistical regression that Massaro and Cohen claimed. Morever, we believe that the superior fit of the additive model to group data demonstrated by Cutting et al (1993) is meaningful.

Individual Data and Noise

Cutting et al (1992) claimed (a) that FLMP absorbs random noise compared to the additive model, and (b) that since individual data are always noisier than group data, FLMP will be advantaged when the two models are run on individual data. Massaro and Cohen (1993) claimed neither is true, but concentrated on the first claim in two ways--addressing the issue of model fits and noise reduction as data accumulates, and addressing the magnitudes of the fits of both models to random data.

Fits of Models to Accumulating Data

As the number of observations per data point increases, noise is reduced. This reduction is guaranteed by statistical regression and the law of large numbers. With this in mind, Massaro and Cohen (1992, Figure 6) presented an analysis of the mean fits of FLMP and the additive model to the data of 21 human subjects in a speech perception experiment. They sampled and pooled the data after 6, 12, and 24 observations per data point. These data are shown in our Figure 1, rescaled to our situation (from RMSD to sum-of-least-squares) and with the abscissa reversed, both for comparison purposes.

Their results showed a clear main effect of accumulated data and also a clear advantage of FLMP for their particular task and design. Critically, however, their data showed no interaction between the fits of the two models and the accumulation of data (reduction of noise). This, they claimed, shows that FLMP does not *differentially* absorb random noise over and above what the additive model absorbs as data become more stable and less noisy. We think this analysis is extremely useful, but we wondered if Massaro and Cohen had sampled the correct range of observations.

--Insert Figure 1 about here--

This accumulative-fit analysis seemed worth replicating on the data of our own 44 subjects. Thus, we fit the two models against the individual data after 1, 2, and 4 observations, and compared those with the fits at the end of the experiment after *n* observations (when *n* was 10, 12, 15, or 60 for each stimulus, typically varying across the 16 stimuli within a data set, depending on the experiment and the number of conditions). These results are also shown Figure 1.

Like Massaro and Cohen's data, our data show a clear main effect of model fits with data accumulation ($F(3,129) = 33.4$, $MSe = 0.307$, $p < 0.001$), and an advantage of FLMP over the additive model ($F(1,43) = 5.25$, $MSe = 0.021$, $p < 0.03$). However, unlike the data of Massaro and Cohen (1993), our data also show an interaction of model fits with data accumulation ($F(3,129) = 5.96$, $MSe = 0.0023$, $p < 0.001$). More pointedly, FLMP bested the additive model on 28 out of 44 individual data comparisons after 1 observation ($p < 0.001$ by a sign test). Notice also this FLMP advantage over the additive model is 64% vs. 36%, quite comparable to the FLMP advantage in random data fits (61% vs. 39%) shown by Cutting et al (1992, Simulation 2). This advantage disappeared after the first observation: FLMP bested the additive model on only 22 of 44 of the individual comparisons after both 2 and 4 observations, and then only 21 of 44 after *n* observations.

Thus, we think Cutting et al (1992) were correct in reporting that FLMP absorbed random error in human data. However, the effect in this task and design manifests itself early in the process of data accumulation. After only two observations in the data sets of Cutting et al (1992), it might appear that the mean fits of FLMP and the additive model were reasonably well determined. However, between 4 and *n* observations the ordinal relation between the fits of the two models reversed for 11 of the 44 subjects. For purposes of modeling, then, we think Massaro and Cohen (1993) are correct to accumulate at least 24 observations per data point to stabilize fits to data.

More importantly, however, we cannot guarantee that the effects of random noise absorption would always stabilize so early in the data accumulation process. Beyond the recommendations of Cutting et al (1992) for research which compares models, we suggest further that researchers plot the comparative fits of their models to data as the data accumulate.

The Relevancy of Fitting Models to Random Data

Massaro and Cohen rightly claimed that the fits of FLMP to random data were highly correlated with those of the additive model. However, they then dismissed as irrelevant the results of simulation to fitting psychological data, since the fits to random data were 8 times worse than the fits to our subject's data. We disagree for reasons stated by Cutting et al (1992): Random data represents all possible data functions, and if one model fits random data better than another model, that advantage ought to be factored out when both models are fit to human or other data of psychological interest.

However, for purposes of argument, let us assume Massaro and Cohen's argument is cogent and correct: "Small RMSDs [root mean squared deviations] only hold for real data, not for random data, and there is not justification for extrapolating from random to real data" (Massaro & Cohen, 1992, ms. p. 16).

New Simulation 3: Fits to Constrained Random Data

What patterns are typical of psychological data? Human data patterns in both Massaro's and our experimental situations are generally logistic and members of the family of psychometric functions. In principle, a psychometric function has a data point at one end of a stimulus continuum higher than the data point at the other end, and all intermediate data points fall in between, usually along a smooth function.

In a third simulation, then, we generated two random numbers between 0.001 and 0.999 and used these values as psychometric endpoints--one for the stimulus with 0 sources of information and the other for the stimulus with 4 sources of information. We then generated 14 different random numbers between these values for each data set--4 for the 1-source stimuli, 6 for the 2-source stimuli, and 4 for the 3-source stimuli.

We generated 600 such data sets, and then fit FLMP and the additive model to them using the NONLIN module of SYSTAT (Wilkinson, 1990) and the sum-of-least-squares method of measuring residuals. Results were similar to those of Simulation 2 in Cutting et al (1992) for fitting unconstrained random data: FLMP fit 353 of the 600 constrained random sets of data (58.8%) better than did the additive model, and the additive model fit 247 better (41.2%), ($F(1,599) = 9.15$, $p < 0.003$).

More interesting, however, are the magnitudes of the fits. The mean sum-of-least-squares fit of both models to these constrained random data sets was 0.178, and the median fit was 0.090. Mean and median fits of the models to the data of the 44 participants of Bruno and Cutting (1988) and Cutting et al (1992) were 0.106 and 0.085, respectively. Since there little difference between these random and human fits, particularly for the medians, Massaro and Cohen's second criticism can be laid to rest: FLMP fits random data better than does the additive model, and does so when the residuals are within in the range of human data. Thus, we maintain, as did Cutting et al (1992), that the differential scope of FLMP over the additive model is something which should be taken into consideration when fitting psychological data of interest.

A Comparison of Directed Perception and Massaro's Paradigm

Multiple Sources of Information

As Massaro and Cohen (1993) noted, there is at least one profound similarity between our two research programs. Massaro's (1987) paradigm and *directed perception* (Cutting 1986, 1991a, 1991b) both ascribe to the ideas that multiple sources of information exist in near-everyday settings for the perception of single objects, events, or their properties; and that perceptual research ought to focus on the use of that multiplicity of information.

This focus is important because perceptual psychologists generally fall into two camps, both outside our domain of methodological commonality. First, many if not most researchers pay little systematic and empirical attention to the idea of multiple sources of information. Research design typically manipulates one source of information at a time, while all other sources are held constant. We and Massaro agree that this methodological choice is a mistake if we are to understand how perception works in everyday contexts. In particular, a one-variable-at-a-time strategy can provide no understanding about the relative importance of various sources of information to a given percept in a given context.

Second, some psychologists oppose in principle the idea of multiple sources of information (e.g. G. Burton & Turvey, 1990). This stance is promoted by a close reading of the works of James J. Gibson on invariants (see Cutting, 1986, 1991a, 1991b). Although we applaud all principled approaches to understanding the concept of information in perception (e.g. Lappin, 1990), we and Massaro agree that this opposition may be a theoretical mistake; we think it ignores the plethora of information so apparent in the world around us.

There are, however, at least two noteworthy differences between directed perception and Massaro's paradigm. These concern the role of combination rules for considering the integration of multiple sources of information, and the role of probability in assessments of information value.

Combination Rules

Directed perception makes no statement about how information is combined. In fact, in some cases combination may not occur at all; some information in a given situation might be selected, other information ignored (Cutting, 1986; Cutting & Millard, 1983). The suggestion and empirical use of an additive combination rule by Bruno and Cutting (1988) was based on its simplicity (Cutting et al, 1992) and on the plausibility of modular systems in brain function (Marr, 1981; Fodor, 1983). While no combination rule is assumed by directed perception; Massaro's paradigm, on the other hand, is explicitly Bayesian and employs the multiplicative combination rule found in FLMP.

Specification and Probability

Directed perception assumes that information specifies what can be perceived. In this way directed perception follows Gibson (1979), but with an important difference, as we will discuss below. Massaro's paradigm, on the other hand, follows Brunswik (1956); it assumes that information is probabalistically related to physical states of affairs.

More concretely, directed perception assumes that the mapping is many-to-1 between information and object/event properties. Given an alert and attending organism, each of several sources of information predicts what could be perceived with a probability of 1.0. In Massaro's paradigm, we would claim, the mapping is many-to-many between information and object/event properties. To make good on this claim, however, we need to explain one aspect of Massaro's paradigm and one logical consequence of it. They are: the imperfect match between information and objects/events, and the residual of that imperfect match.

First, Massaro's paradigm assumes each source of information predicts what could be perceived with a probability less than 1.0. That is, following Brunswik (1955, p. 207), Massaro believes there is a "limited ecological validity or trustworthiness of cues." Indeed, the fuzzy-logical model of perception (FLMP) explicitly assumes that no information-object relation can have a probability of 1.0. Inspection of the FLMP equation (see, for example, Cutting et al, 1992, Equation 2) shows that the use of two unit probabilities—one source of information present (weighted 1.0) and one source absent (weighted 0.0)—will place a zero in the denominator of the equation and create an indeterminate solution. Thus, by methodological design, sources of information must be imperfectly matched to objects and events.

Second, if one source of information does not fully predict a particular object or event, it must then also predict at some residual value yet another object or event. To be concrete, FLMP assumes that no source of information can predict a particular object or event with a probability of 1.0. Thus, if Source A predicts Object 1 for 95% of all its occurrences, then it for the other 5% of its occurrences it must predict something else, say Object 2 (or, perhaps more likely, a collection of Objects 2 through *n*). Thus, the fact of imperfect matches between a source and an object implies a residual of occurrences between that source and other objects or events. These two facts—nonunitary information/object probabilities and their residuals—make Massaro's paradigm an explicit example of *indirect perception*, as outlined by Cutting (1986).

Three Theories of Perceptual Information

There are at least three perspectives on the issue of mapping between information and objects. These are not theories of perception but are logically prior to any formulation of a perceptual theory; they are theories of how perceptual information is related to the world. In the case of vision, as shown in Figure 2, they are structural representations of the problem of inverse optics, or mapping backwards from the two dimensional

(2D) image surface to the three dimensional (3D) world. In particular, notice that there is no perceiver in any of the three mapping diagrams; everything shown is logically prior to information registration at a perceiver's sense organ.

--Insert Figure 2 about here--

The most straightforward theory is Gibson's (1979) *direct perception* as shown in the left panel; it assumes 1-to-1 mappings between information and the objects/events specified (G. Burton & Turvey, 1990; Cutting, 1991a). Indeed, the notion of a single invariant available to be picked up during perception is exactly such a specification. Under direct perception the perceived world and its perceptual information are perfectly matched.

Slightly more complex in a logical sense is *directed perception* (Cutting, 1986), shown in the middle panel. It assumes there can be more than one invariant, or more than one source of information, that can specify what is to be perceived. Note, however, that the number of possible information sources necessarily outstrips the number objects or events to be perceived, because no source of information is related to more than one object or event. Under directed perception information is a rich beyond the world of perceived objects and events; but information is orderly and ordered.

Considerably more complex is *indirect perception*, shown in the right panel. It was long-criticized by Gibson, and most clearly instantiated in the work of Brunswik (1956) and now in that of Massaro (1987). It assumes, according to the logic above, a many-to-many mapping between information and objects/events. Under indirect perception information is rich, but there are now tangles of mappings with the perceived world to be unraveled by probabalistic weights and the psychological processes they imply. The major problem inherent in the plexus of associations denoted by indirect perception is perhaps best implied by the title of Gibson's (1957) review of Brunswik's book (1956): "Survival in a world of probable objects:" Evolution and natural selection have a difficult enough time without having individuals and species working against a shifting array of untrustworthy perceptual information.

Clearly, the word liable to create most mischief in this analysis is *specification*. Previous discussion of the concept (Cutting, 1986, 1991a, 1991b; see also Gibson, 1979; Michaels & Carello, 1981) has been incomplete and thus apt to create confusion. Massaro and Cohen (1993) rightly noted its incompleteness.

What Does Depth Information Specify About Depth?

Our notion of specification in this context entails two ideas. First, specification of the relative depth of objects does not assume a particular metric. Second, whatever metric is specified, that specification does not come free; it must be based on a set of assumptions about how 2D image properties map back onto the 3D objects in the world. These assumptions are deeply related to what are sometimes called "nonaccidental properties" (Witkin & Tenenbaum, 1983), discussed with respect to the problem of inverse optics as studied in the field of machine vision. Let us consider the issue of measurement specification first, then then the combined issue of measurement and its underlying assumptions.

Starting with Stevens (1951), some depth specifications might have ordinal properties, some interval properties, some perhaps even ratio properities. Such different specifications are, for example, inherent in the coding and modeling of information and depth relations found in the work of Maloney and Landy (1989; Landy, Maloney, and Young, 1991), although their conclusions are different than what we present here. For purposes of continuity with our previous research it will be useful for us to consider the four sources of information about depth investigated by Bruno and Cutting (1988; Cutting et al 1992)--relative size, height in plane, occlusion, and motion parallax.

Ordinal Specification

Without additional assumptions or knowledge on the part of the observer, most sources of information about depth are probably ordinal. For example, the occlusion of one object by another specifies only depth order--which object is closer to the observer. Although occlusion offers only this weak measurement, it is a very strong source of visual information about depth because it is based on three reasonable assumptions, two from Euclid's optics (H. Burton, 1945) and one from Gestalt psychology (e.g. Koffka, 1935)—the linearity of projecting rays, the general opacity of objects, and the good continuation of boundary contours. With respect to the latter, the boundary contour of an occluding object will likely possess good continuation everywhere; the boundary contour of the occluded object, on the other hand, will typically have abrupt discontinuities. The existence of discontinuities specifies occlusion, which in turn specifies depth order.

Height in the projection plane (or projection surface) also typically provides only ordinal information, as shown in Figure 3. That is, the reverse mapping from the 2D image to the 3D world promises only ordinal increase in depth, but it promises that quite firmly. That promise is based on three assumptions: linearity of projecting rays, the presence of gravity (thus, objects will rest on the surface of support), and the opacity of the support surface. If one goes further, beyond what is shown in Figure 3, and assume a *plane* of support (which could be specified by other information) and a familiarity (knowledge) of one's eye height, height in plane has the potential for offering much more than ordinal information. These assumptions, however, can occasionally be costly and constraining. Few natural surfaces of support are planar beyond a local neighborhood, playing havoc with the global assumption of a constant and familiar eye height, and with any bets about intervals or ratios of distances. Thus, it seems best to conclude that height in plane, like occlusion, provides only ordinal information.

--Insert Figure 3 about here--

Unanchored Ratio Specification

As suggested above, some sources of information have the potential for offering more than ordinality. Relative size is a good example. In many environments one can assume that similarly appearing objects or textures are the same physical size, or deviating stochastically about some mean. Books, bushes, cars, cattle, grass, gravel, houses, litter, people, and trees all demonstrate this property. In these cases and in others the relative sizes of objects as projected on the retina have the potential for offering powerful information: Projective size varies inversely and linearly with distance. Thus, if two same-shaped objects appear in the optic array and one is half the size of the other, then it is also twice the distance. However, if one does not have knowledge about the identity, and thus the size, of the particular object in question, this information is unanchored and confounded by a scale factor: The two objects could be at 1 and 2 m, or at 100 and 200 m. In other words, information in relative size scales depth, but by itself and without prior knowledge provides no interval information to scale its ratios.

Assuming the rigidity of objects in the world, the information in motion parallax may be similar to that in relative size. However, if one has knowledge of the extent of one's movement, motion parallax has the potential of providing absolute information about depth (Maloney & Landy, 1989). One could only gain this knowledge, however, through nonvisual (proprioceptive) means, knowledge of eye height, or knowledge of the size of the objects which demonstrate the parallax due to observer movement. Thus, we believe the parallactic motion of a near and a far panel against a central panel by itself only scales depth, and provides information about the relative, but not absolute spacing, of objects.

An Analogy

In these four cases—occlusion, height in plane, relative size, and motion parallax--and in others, depth relations are specified, but in none of these cases is absolute, metric depth determined. Thus, observers judgments about exocentric depth—the dependent measure in the direct scaling studies of Bruno and Cutting (1988) and of Cutting et al (1992)—are ordinally constrained along some stimulus dimensions, and perhaps ratio constrained on others. Understanding how viewers rally their use of a subjective depth scale is important in trying to understand how they perceive the distance relationships of objects in space. How do they do this?

We are most interested in how nonmetric sources of information might combine to give accurate, near-metric information about depth. This idea is perhaps best captured by an analogy: Just as the multidimensional scaling of nonmetric information can, through multiple constraints on ordinality, converge on a metric representation (e.g. Shepard, 1980), so too the integration and combination of nonmetric (ordinal) sources of information about depth may also converge on a metric representation of the spatial layout of objects around us. We find such a possibility likely and fascinating, although the mechanisms by which it might occur are currently unclear and any understanding must be left for further research.

Conclusions

In conclusion, we partly agree with Massaro and Cohen (1993) on several points, but we disagree on others. First, we agree that modeling analyses on group data must proceed carefully, but we think Massaro and Cohen's criticisms are not pertinent to our case. Second, we agree that Massaro's paradigm and directed perception share at least one important property—they embrace the idea of multiple sources of information. However, they can be distinguished in two critical ways: the choice and importance of a combination rule, and the theory of perceptual information they endorse. Third, we agree that our concept of information specification was previously ill-defined, and we have tried to define it here more precisely with respect to inherent measurement scales and assumptions underling those scales.

We disagree with Massaro and Cohen (1993) on several points. First, they claimed that FLMP can be prudently fit to group data but that the additive model cannot. However, we demonstrated that both models can fit group data better when all the individual data are better fit by the other model. Thus, group analyses by FLMP are on the same logical footing as those by the additive model. We think the fits of the models to group data become meaningful under circumstances where both models fit individual data about equally well. Second, we disagree with Massaro and Cohen's (1993) belief that comparison of the fits of models to random data are irrelevant to their fits to psychological data. Moreover, we countered their specific claim that the fits to random data should be dismissed on grounds of their being "unacceptably large" (ms p.16). We demonstrated that when random data are constrained to fit within the bounds of endpoints of a response continuum, the fits of both FLMP and the additive model are quite close to the fits of the models to psychological data, and that FLMP maintains about the same advantage over the additive model. This result reinforces the statement of Cutting et al (1992) that the relative ability of a model to fit random data should be neutralized when fitting psychological data of interest. It also suggests a new caveat: the fits of models to individual data should be compared over the course of data accumulation to insure that differential fits of models are not due to the absorption of random noise.

References

Bruno, N. & Cutting, J.E. (1988). Minimodularity and the perception of layout. *Journal of Experimental Psychology: General, 117*, 161-170.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193-217.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California.

Burton, G. & Turvey, M.T. (1990). Perceiving the lengths of rods that are held but not wielded. *Ecological Psychology, 2,* 295-324.

Burton, H.E. (trans.) (1945). The optics of Euclid. *Journal of the Optical Society of America, 35*, 357-372.

Cutting, J.E. (1982). Two ecological perspectives: Gibson vs. Shaw and Turvey. *American Journal of Psychology, 95,* 199-222.

Cutting, J.E. (1986). *Perception with an eye for motion*. Cambridge, Mass.: MIT Press/Bradford Books.

Cutting, J.E. (1991a). Four ways to reject directed perception. *Ecological Psychology, 3,* 25-34.

Cutting, J.E. (1991b). Why our stimuli look as they do. In G. Lockhead and J.R. Pomerantz (eds.*), Perception of structure: Essays in honor of Wendell R. Garner* (pp. 41-52). Washington, D.C.: American Psychological Association.

Cutting, J.E, Bruno, N., Brady, N.P., & Moore, C. (1992). Selectivity, scope, and simplicity of models: A lesson from fitting judgments of perceived depth. *Journal of Experimental Psychology: General, 121*, 362-381.

Cutting, J.E. & Millard, R.T. (1983). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology: General, 113,* 198-216.

Fodor, J.A. (1983). *Modularity of mind.* Cambridge, Ma: MIT Press/Bradford Books.

Gibson, J.J. (1957). Survival in a world of probable objects. *Contemporary Psychology, 2,* 33-35.

Gibson, J.J. (1979). *The ecological theory of visual perception*. Boston: Houghton Mifflin

Koffka, K. (1935). **Principles of Gestalt psychology**. New York: Harcourt

Landy, M.S., Maloney, L.T., & Young, M.J. (1991). Psychophysical estimation of the human depth combination rule. In P.S. Shenker (ed.) *Sensor fusion III: 3-D perception and recognition, Proceedings of the SPIE, 1383,* 247-254.
Lappin, J.S. (1990). Perceiving the metric structure of environmental objects from motion, self-motion, and stereopsis. In R. Warren & A.H. Wertheim (eds.) *Perception & control of self-motion* (pp. 542-579). Hillsdale, N.J.: Erlbaum.
Maloney, L.T. & Landy, M.S. (1989). A statistical framework for robust fusion of depth information. In W.A. Perlman (ed.) *Visual communication and image processing, IV, Proceedings of the SPIE, 1199,* 1154-1163
Marr, D. (1981). *Vision*. San Francisco: Freeman
Massaro, D.W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, N.J.: Erlbaum.
Massaro, D.W. (1988). Ambiguity in perception and experimentation. *Journal of Experimental Psychology: General, 117,* 417-421.
Massaro, D.W. & Cohen, M.M. (1993). The paradigm and the fuzzy logical model of perception are alive and well. *Journal of Experimental Psychology: General, 122,* 115-124.
Michaels, C.F. & Carello, C. (1981). *Direct perception*. Englewood Cliffs, N.J.: Prentice-Hall.
Shepard, R.N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science, 210,* 390-398.
Stevens, S.S. (1951). Mathematics, measurement, and psychophysics. In S.S. Stevens (ed.*) Handbook of experimental psychology* (pp. 1-49). New York: Wiley.
Wilkinson, L. (1990). *SYSTAT: The system for statistics*. Version 5.2. Evanston, IL: SYSTAT, Inc.
Witkin, A.P. & Tenenbaum, J.M. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld (eds.) *Human and machine vision* (pp. 481-543). New York: Academic.

Footnotes

[1] Cutting et al (1992) found in their experiments (a) that the fits to the 34 individual data sets showed no difference between the models (18 favored the additive model, 16 favored FLMP), with a slight mean advantage for FLMP ($t(33) = 1.142$, $p > 0.26$); (b) that the group modeling data showed a 3-to-1 advantage in residuals for the additive model (0.0091 vs. 0.0271); and (c) that the analysis of variance on difference scores in the judgment data showed a subadditive trend, a result favoring FLMP. Thus, the reported inconsistencies were: (a) was indeterminant, (b) favored the additive model, and (c) favored FLMP. Massaro and Cohen (1992) first dismissed (b), our group modeling results, which for reasons given in a later section we think is based on a false assumption about our analysis. Second, they relied on our report of a *t* test (FLMP slightly favored, $p > 0.26$) in (a), rather than on the fact that data of less than half the subjects' data were better fit by FLMP. We find their choice of evidence charming since it is an analysis of *group* data, not of individuals (where FLMP was favored for only 16 of 34 individuals). However, even if one dismissed (b), the group modeling results, we think the inconsistency between (a) and (c) remains.

[2] Massaro and Cohen (1992) misread our claim about FLMP and scope. Nowhere did we claim FLMP was "superpowerful" and unfalsifiable; our claim was only that FLMP has greater scope and is therefore less falsifiable than other models, particularly the additive model. However, we are pleased that Massaro and Cohen chose to fit FLMP and other models to random data; this was precisely our recommendation for further research within Massaro's paradigm, and is the prescription (Rx) noted in our running head for it's continued health.

[3] This is the only direct response by Massaro and Cohen (1992) to the conclusions of Cutting et al (1992), but it contains only an admonition and a diversion. First, against the relation of equation length to a model's ability to fit human data, Massaro and Cohen warned against imputing causation from correlation. We think Cutting et al were properly cautious in their statements. Second, Massaro and Cohen noted that if the input data to the models were *z* transformed the equation for FLMP would then be simpler than that for the additive model. We don't understand the force of this argument; it trades measuring simplicity of operations *within* the models for measuring the simplicity of operations *on* raw data. The additive model needs no data transformation for its inputs, and thus is simpler here as well. Massaro and Cohen's admonition and diversion aside, we maintain that, given our data and algorithmic information theory, equation length is not unreasonable as a consideration when comparing models.

Table 1

Residual sum of least squares for the additive model and FLMP in their fits to the data of simulated individual subjects and groups. Simulation 1 replicates Massaro and Cohen's (1992) simulation; Simulation 2 shows exactly the opposite result and impugns their conclusions about the efficacy of fitting group data with FLMP.

	Set Performance Level					Fits to Models
	Number of Sources of Information					
Simulation 1	0	1	2	3	4	FLMP additive
1	.200	.900	.999	.999	.999	.000 < .340
2	.000	.050	.400	.850	.999	.007 < .107
3	.000	.000	.000	.050	.100	.002 < .004
4	.400	.400	.500	.600	.600	.009 < .010
Mean of Individuals						.005 < .115
Group Mean	.150	.338	.475	.625	.675	.009 > .008
Simulation 2						
1	.333	.700	.700	.700	.999	.115 > .114
2	.600	.800	.800	.800	.999	.040 > .039
3	.000	.300	.600	.750	.950	.032 > .031
4	.600	.700	.800	.900	.950	.003 > .002
Mean of Individuals						.048 > .047
Group Mean	.383	.625	.725	.787	.975	.020 < .021

Note: These data represent 16 stimuli, 1 with 0 sources of information, 4 with 1 source, 6 with 2 sources, 4 with 3 sources, and 1 with 4 sources, with no variability at any level.

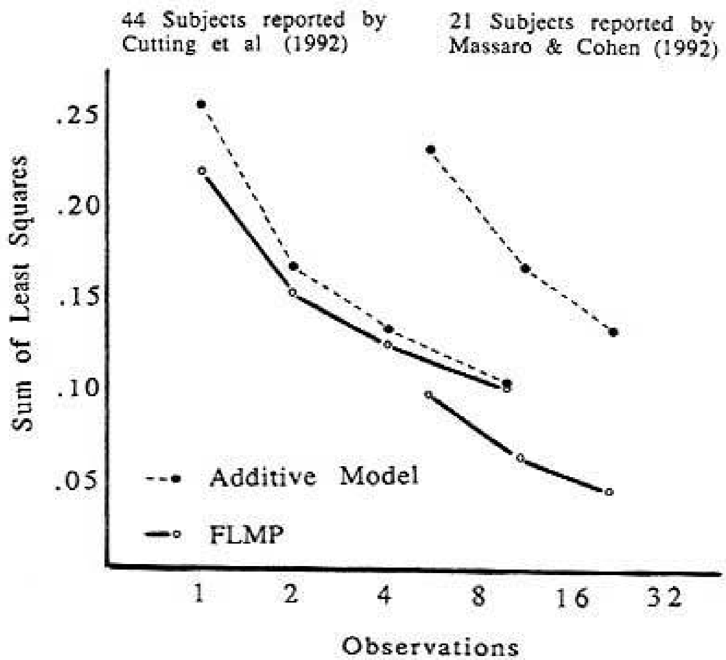

Figure 1: The relative fits of FLMP and the additive model to the data of 21 individuals reported by Massaro and Cohen (1992) and the 44 individuals reported by Cutting et al (1992), as a function of how they accumulate with more observations per data point. (The data of Massaro and Cohen are plotted after 6, 12, and 24 observations; those of Cutting et al are plotted after 1, 2, 4, and *n* observations, where n is between 10 and 60. Most importantly, the data of Massaro and Cohen show no reliable interaction of model fits with increasing number of observations, whereas the data of Cutting et al show such an interaction.)

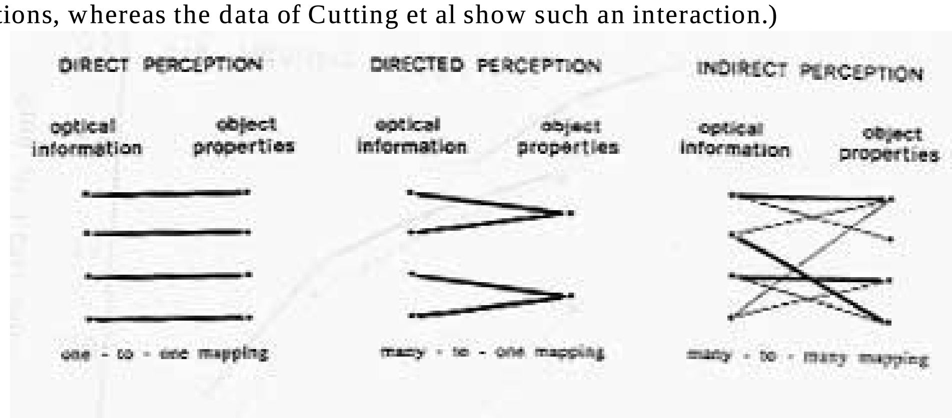

Figure 2: Three theories about the mapping between visual information and the visual objects, events, and their properties which they represent. (The left panel shows the representation for Gibson's (1979) direct perception, where each source singly specifies the object or object property to be seen. Mappings are one-to-one. The middle panel shows Cutting's (1986) directed perception, where sources multiply specify objects or their properties. Mappings from information to object (or object property) are many-to-one. In direct and directed perception the probability of the information specifying the object property is 1.0. The right panel shows the probabilistics mappings of indirect perception as exemplified by Brunswik (1956) and Massaro (1987). Here, mappings are many-to-many, with weights indicated by thickness of lines. Adapted from Cutting, 1986; see also Cutting, 1991b.)

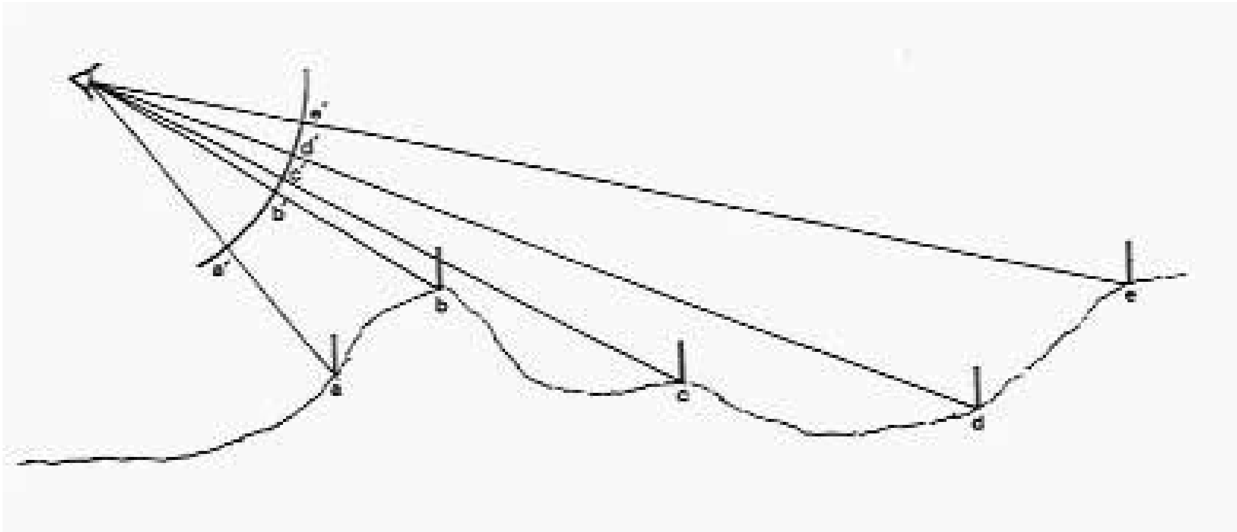

Figure 3: The ordinal depths relations specified by the relative height in plane of the base of objects. (Only linearity of projective rays, object opacity, and gravity need be assumed. Planarity of the ground plane and knowledge of eye height is not assumed here. From Cutting, 1986).