5) *Top halves of compound eyes covered.* Bees with their ocelli and the bottoms of their compound eyes functioning were flight-tested in daylight and at dusk. All could maintain level flight during both periods, and treated members of both colonies were able to spiral and beeline whenever normal bees were doing so, until intensities fell to 0.3 to 0.5 lux. Treated bees from the nearby colony also were able to zigzag homeward by landmarks whenever their normal counterparts did so.

The results show that the western bumblebee can use its dorsal ocelli alone or in conjunction with the tops of its compound eyes to steer by polarized light. The compound eyes are necessary for maintaining altitude, steering by landmarks, and alighting: that is, for any task requiring perception of form or color. But the larger-apertured ocelli appear to function longer than the small-faceted compound eyes at dusk. With ocelli, the bees can prolong their foraging by continuing to use the directional pattern overhead when the surroundings are too dimly lit for their compound eyes to distinguish landmarks.

Perception of polarized light is not restricted to the ocelli, even in the bumblebee. There are many anocellate species which can perceive polarized light, whether or not they utilize it. An extreme example, the European earwig, *Forficula auricularia* L., is anocellate, photonegative, and highly nocturnal, yet it responds weakly by day and more strongly at dusk to rotation of the Polaroid (*5*). And many other kinds of anocellate insects have enlarged facets in the upper parts of their compound eyes, in effect combining ocellar and ommatidial functions. So dorsal ocelli are no better correlated with polarized-light navigation than they are with the presence of wings or the diurnal habit (*12*).

There are, in fact, so many types of dorsal ocelli that it would be unrealistic to expect only one function (*13*). But conflicting roles have been ascribed to them (*14*), even though most texts still list them with other "stimulatory organs," a group for which no satisfactory explanation exists. There might be less confusion if insects were observed more often in their natural surroundings, where an ocellar linkage with behavior should be easier to discern than it has been indoors (*6*, *7*). When the habits of an insect in its natural setting include directed travel by daylight or twilight, steering by zenith polarization patterns becomes one ocellar function worth considering.

W. G. WELLINGTON
*Department of Plant Science* and
*Institute of Animal Resource Ecology,*
*University of British Columbia,*
*Vancouver, Canada, V6T 1W5*

### References and Notes

1. M. Lindauer and B. Schricker, *Biol. Zentralbl.* **82**, 721 (1963); B. Schricker, *Z. Vgl. Physiol.* **49**, 420 (1965).
2. M. Lindauer, *Annu. Rev. Entomol.* **12**, 439 (1967). Some greens appear gray to the honeybee eye even in daylight. At dusk, large parts of that gray scale would be virtually indistinguishable.
3. K. von Frisch, *Bees: Their Vision, Chemical Senses, and Language* (Cornell Univ. Press, Ithaca, N.Y., 1971), pp. 113–136. As the greatest percentage of polarized light lies midway between the solar and antisolar points, the band of most intense polarization near sunrise and sunset runs north-south through the zenith.
4. Hymenoptera: Apidae. H. E. Milliron, *Mem. Entomol. Soc. Can. 82* (1971), pp. 58–67.
5. This is part of a study of diurnal changes in the ability of several kinds of flying insects to steer by plane-polarized light (W. G. Wellington, in preparation).
6. Overhead polarization failed regularly around midday near summer solstice, when unpolarized glare near the high sun invaded the zenith, and more sporadically whenever sufficiently heavy clouds, smoke, or haze obscured the zenith. The increasingly smoggy atmosphere near major industrial areas is making outdoor tests more difficult.
7. W. G. Wellington, *Nature (Lond.)* **172**, 1177 (1953); *Ann. Entomol. Soc. Am.* **48**, 67 (1955).
8. Each ground test involved a pair of treated and normal bees. As flight tests were synchronized with types of normal flight or particular states of the sky, each treated bee could be compared with four to eight normal workers flying simultaneously. The following data refer only to bees treated. Group 1, 20 on the ground, 30 in the air; all made multiple flights. Group 2, 26 of 40 on the ground crawled, 50 of 72 in the air flew, none returned. Group 3, 30 flew, none were recovered. Group 4, 30 on the ground; 50 from the nearby colony in the air, all flew, all recovered; 40 from the distant colony could not leave feeding site. Group 5, 50 tested from each of the two colonies; 5 were recovered from that nearby, none from the distant colony.
9. A Kahlsico light meter, model 268WA620, was read in the center of a lawn 24 m in diameter with a sensor 1 m above the ground, facing northward, not skyward.
10. F. Schremmer, *Oesterr. Zool. Z.* **2**, 242 (1949).
11. K. von Frisch, M. Lindauer, K. Daumer, *Experientia* **16**, 289 (1960).
12. H. Kalmus, *Proc. R. Entomol. Soc. Lond. Ser. A Gen. Entomol.* **20**, 84 (1945).
13. L. J. Goodman, *Adv. Insect Physiol.* **7**, 97 (1970); A. Wolsky, *Biol. Rev.* **8**, 370 (1933).
14. H. Autrum and N. Metschl, *Z. Vgl. Physiol.* **47**, 256 (1963); E. Bozler, *ibid.* **3**, 145 (1925); H. Homann, *ibid.* **1**, 541 (1924); S. H. Chen and B. Young, *Sinensia* **14**, 55 (1943); D. A. Parry, *J. Exp. Biol.* **24**, 211 (1947).
15. I thank the Polaroid Corporation for the sheet of Polaroid. Supported by grant A 6239 from the National Research Council of Canada.

2 November 1973

## Testing of Computer-Assisted Methods for Classification of Pharmacological Activity

Recently Ting, Lee, Milne, Shapiro, and Guarino (*1*) have investigated the relationship between the pharmacological activity of a set of 66 sedatives and tranquilizers and their mass spectra. They applied several distance criteria for judging mass spectra as similar or different. With the aid of a computer they could transform the mass-spectral data so that drugs with similar pharmacological activity would be "close," and those with different pharmacological activity would be "distant." They then found that they could distinguish sedatives from tranquilizers and thus classify test compounds on the basis of their "neighbors."

It seemed to me quite likely that the authors' set of drugs is not suitable for testing the relationship between pharmacological activity and mass spectrum. The sample population lacks independence, inasmuch as some drugs are obviously related to others. More than half the sedatives are barbiturates and more than half the tranquilizers are phenothiazines. Thus it is not surprising that Ting *et al.* (*1*) could classify cyclobarbital with the other barbiturates. With some of their other successes dependence is not so obvious, and some may be nontrivial. Yet it is possible that a large portion of the set of sedatives (or of tranquilizers) is made up of several families of compounds with similar structures and mass spectra. If so, it is trivial that the mass-spectral data can be transformed so that similar drugs cluster. And the ability to classify a drug would merely reflect the fact that it is a member of a family of sedatives (or tranquilizers). Thus it is not clear to what extent the authors' success is due merely to this lack of independence in the sample population.

To assess the independence, I have applied the authors' simplest nearest-neighbor method just to the *names* of the drugs. To the extent that the drugs are clustered into families, there will be drugs of similar pharmacological activity with similar names. And a test drug may be classified according to the pharmacological activity of drugs with

"neighboring" names. The names may be considered as points in a 26-dimensional space. The first component is the number of A's in the name, the second is the number of B's, and so forth. Given names $N = (n_A, n_B, \ldots, n_Z)$ and $M = (m_A, m_B, \ldots, m_Z)$, we may define the distance between them by

$$d_{N,M} = [(n_A - m_A)^2 + (n_B - m_B)^2 + \ldots + (n_Z - m_Z)^2]^{1/2}$$

If $d_{N,M}$ is small, then the names $N$ and $M$ are close. This distance measure is quite simple and easy to apply, although it lacks scaling (*2*) and it loses the information contained in the ordering of the letters (whereby names ending in "barbital" are obviously similar). To classify a test drug, the nearest-neighbor method was then used: a drug was classified as a sedative or a tranquilizer according to whether the drug closest to it was a sedative or a tranquilizer. Clearly, this is not proposed as a useful method for classifying drugs, but rather as a convenient one for testing the independence of the sample. If the sampling of drugs were truly independent, one would expect a classification method based simply on names to succeed only 50 percent of the time. Yet when this method was applied to each of the 64 different names (*3*), 54 of them (84 percent) were classified correctly (*4, 5*). For comparison, Ting *et al.* (*1*), applying this nearest-neighbor method to the mass spectra, classified 83 percent correctly.

Thus I conclude that the set of drugs contains too many obviously related compounds, and that this set is therefore not adequate to demonstrate a relationship between pharmacological activity and mass spectrum. Such a relationship may exist, but Ting *et al.* (*1*) have failed to demonstrate it. They have demonstrated only the tautology that similar compounds have similar properties.

This reflects quite a general problem in *testing* any method for classifying drugs. A method that has been proved capable of distinguishing sedatives from tranquilizers among one set of drugs may fail when presented with a new structure. An adequate test of a method would require a set of drugs that are independent. Otherwise, whatever success is observed might be due in part merely to the method's ability to recognize dependence. Yet independence is difficult to achieve, or even to test for. In developing new drugs, chemists often rely upon analogy with known ones, so that it is unlikely that any set of drugs is sufficiently independent to permit a thorough test of a classification method.

CHARLES L. PERRIN*
*Institute of Neurobiology, Göteborg University, Göteborg, Sweden*

### References and Notes

1. K.-L. H. Ting, R. C. T. Lee, G. W. A. Milne, M. Shapiro, A. M. Guarino, *Science* **180**, 417 (1973).
2. B. R. Kowalski and C. F. Bender, *J. Am. Chem. Soc.* **94**, 5632 (1972).
3. Among the 66 drugs there are two pairs of positional isomers. Of course, a distance approach would also consider names of positional isomers as quite similar.
4. Ethchlorvynol, *ethinamate*, glutethimide, *captodiamine*, *methapyrilene*, *pyrilamine*, meprobamate, *tybamate*, *ectylurea*, and chloprothixene were assigned incorrectly. Both oxanamide and methdilazide had as two equidistant nearest neighbors one sedative and one tranquilizer, but the three next nearest neighbors were all tranquilizers.
5. A test compound was also classified as a sedative or tranquilizer according to whether a majority of its three nearest neighbors were sedatives or tranquilizers. By this method only six drugs were assigned incorrectly—those italicized in (*4*). Presumably, appropriate refinements, such as weighting of data, would reduce the error further, as in (*1*).
6. I thank Professor Holger Hydén, Institute of Neurobiology, Göteborg, for his hospitality and the Institute of Physiology, Göteborg, for the use of a Hewlett-Packard 9830A desk calculator.
* Special research fellow of the NINDS while on sabbatical leave (1972–73) from the department of chemistry, University of California, San Diego, La Jolla 92037, to which I have returned since writing this comment.

2 July 1973

Perrin's comment that our training set of drugs lacks independence is valid. In our report (*1*), we stated that "Half the sedatives are in fact barbituric acid derivatives. . . . That any of these can be correctly classified comes as no surprise. . . ." Our whole point is that other structurally diverse types also correlate, as Perrin admits. We plan to discuss in detail the applicability of the methods to different situations (*2*). It is not particularly surprising that the letters of the trivial names correlate as well, since many reflect structural features of the molecule (-ital, -azine). It is, in fact, usually much easier to group a series of compounds by their structures (or less obviously by their trivial names) than by their mass spectra, since the latter are often grossly changed by relatively small changes in substituent (for example, the phenothiazines).

The structures of drugs have often been correlated with activity. However, predicting chemical reactivity from structure alone is very difficult. Another way of gaining information on reactivity is to subject a drug to physical techniques such as mass spectrometry, nuclear magnetic resonance, infrared spectroscopy, polarography, and so forth, and observe whether or not the experimental results themselves correlate directly with activity. In the case of mass spectrometry, the molecule is subject to ionization with high-energy electrons, which results in extensive bond breaking and also bond formation. Therefore, the cracking pattern depends heavily on the molecular structure, sometimes even in regard to geometrical aspects (isomers). So long as these experimental results correlate with pharmacological activity, it is not essential that they result from a physical process which closely simulates body reactions. Indeed, spectral results may offer a different view of reactivity than that obtained from consideration of the structure alone. We may even expect some insight (*3*) concerning the nature of the pharmacological actions. Also, it is conceivable that a common mass spectral pattern correlating with sedative activity may be found in a group of structurally diverse compounds, because under electron bombardment they may all go, in some small part, to a common intermediate.

KAI-LI H. TING
*Division of Computer Research and Technology, National Institutes of Health, Bethesda, Maryland 20014*

### References and Notes

1. K-L. H. Ting, R. C. T. Lee, G. W. A. Milne, M. Shapiro, A. M. Guarino, *Science* **180**, 417 (1973).
2. K-L. H. Ting, R. C. T. Lee, C. L. Chang, A. M. Guarino, in preparation.
3. We have recently found (*2*) that a compound that shows a change in activity due to a slight change in structure can be detected from its mass spectrum by using the Fisher discriminant and feature selection methods.
4. I am grateful to H. M. Fales, M. Shapiro, G. W. A. Milne, A. M. Guarino, and R. C. T. Lee for their helpful discussions and assistance regarding this reply.

12 September 1973; revised 23 October 1973