(i) Unit responses to electromagnetic FCU stretch closely resembled in form and latency their responses to extension displacement, which also stretched FCU. (ii) Studies in awake relaxed animals indicate that few units in area 4 respond to stimulation similar to that delivered to other tissues by FCU movement (*6*, *7*).

The data are in accord with results in acute anesthetized preparations (*8*, *9*) and awake relaxed animals (*6*) indicating that short-latency responses to muscle stretch occur in primary motor as well as primary sensory cortex. The predominance of phasic responses and the small magnitude of FCU stretch ($< 75$ $\mu$m in the cadaver) suggest that group 1A muscle afferents play a significant role. The large number of units giving both ON and OFF responses, also noted by Hore *et al.* (*9*), is in contrast to the behavior of the peripheral stretch receptors (*10*).

The most striking aspect of the data is the high proportion, in all three cortical areas, of displacement-responsive units that also responded strongly to FCU stretch. This result is particularly impressive since FCU is only one of 11 muscles (six flexors and five extensors) involved in wrist flexion and extension (*11*), and the FCU stretch by the slug ($< 75$ $\mu$m in the cadaver) was less than that by the extension displacement (100 to 200 $\mu$m). Combined with the similarity in latency and form between FCU stretch responses and extension displacement responses, the result implies that muscle stretch is a major factor in the short-latency response of area-4 units to limb displacement and suggests that such stimulation has a prominent role in motor control at the cortical level. At the same time, the high intensity of the cortical unit response to the stretch of a single flexor muscle implies that the motor cortex response is not proportional to the number of receptors stimulated. Thus the data do not support the hypothesis of a graded transcortical servo loop (*12*), unless one or more additional assumptions are made. One possibility is that the inputs from synergist muscles act in parallel, so that the stretch of any one is equivalent on the cortical level to the stretch of all. Another possibility, supported by units such as the one in Fig. 2C, is that movement of joints and other tissues, which was marked with the displacements, can inhibit the cortical response to muscle stretch.

JONATHAN R. WOLPAW
*Laboratory of Neurophysiology, National Institute of Mental Health, Bethesda, Maryland 20014*

### References and Notes

1. E. V. Evarts, *Science* **179**, 501 (1973); B. Conrad, K. Matsunami, J. Meyer-Lohman, M. Wiesendanger, V. B. Brooks, *Brain Res.* **71**, 507 (1974); R. Porter and P. M. H. Rack, *J. Physiol. (London)* **241**, 95P (1978).
2. J. R. Wolpaw and T. R. Colburn, *Brain Res.* **141**, 193 (1978); T. R. Colburn, J. R. Wolpaw, W. Vaughn, in preparation.
3. E. V. Evarts, in *Methods in Medical Research*, R. F. Rushmer, Ed. (Year Book, Chicago, 1966), vol. 2, p. 241.
4. Histologic criteria of T. S. Powell and V. B. Mountcastle [*Bull. Johns Hopkins Hosp.* **105**, 108 (1959)].
5. P. Grigg and B. J. Greenspan, *J. Neurophysiol.* **40**, 1 (1977).
6. R. N. Lemon and R. Porter, *Proc. R. Soc. London Ser. B* **194**, 313 (1976).
7. Y. C. Wong, H. C. Kwan, W. A. Mackay, J. T. Murphy, *J. Neurophysiol.* **41**, 1107 (1978).
8. M. Wiesendanger, *J. Physiol. (London)* **228**, 203 (1973); G. E. Lucier, D. C. Ruegg, M. Wiesendanger, *ibid.* **251**, 833 (1975).
9. J. Hore, J. B. Preston, R. G. Durkovic, P. D. Cheney, *J. Neurophysiol.* **39**, 484 (1976).
10. P. B. C. Matthews, *Mammalian Muscle Receptors and Their Central Actions* (Williams & Wilkins, Baltimore, 1972), pp. 140–187. Since termination of electromagnetic FCU stretch presumably phasically stretched the small portion of muscle (15 percent of the muscle bulk) distal to the slug as well as the distal tendon, it is possible that some of the apparent OFF responses were actually ON responses resulting from input from muscle spindles and Golgi tendon organs in these distal regions.
11. H. O. Kendall, F. P. Kendall, G. E. Wadsworth, *Muscle Testing and Function* (Williams & Wilkins, Baltimore, 1971).
12. T. H. Koeze, C. G. Phillips, J. D. Sheridan, *J. Physiol. (London)* **195**, 419 (1968); C. G. Phillips, *Proc. R. Soc. London Ser. B* **173**, 141 (1969).
13. I thank Dr. E. V. Evarts, Dr. J. E. Reed, and Mr. A. C. Ziminsky for invaluable advice and assistance.

26 June 1978; revised 22 November 1978

## To Know with the Nose: Keys to Odor Identification

Abstract. *Successful odor identification depends on (i) commonly encountered substances, (ii) a long-standing connection between an odor and its name, and (iii) aid in recalling the name. The absence of any one ingredient impairs performance dramatically, but the presence of all three permits ready identification of scores of substances, with performance seemingly limited only by the inherent confusability of the stimuli.*

How many common substances can a person identify by smell? Estimates have varied from about 6 to 22 when subjects have had a single chance to identify each substance (*1*-*5*). For instance, 200 persons (physicians, nurses, medical students, and patients with normal olfaction) could identify an average of only 6 out of 12 odorants (*1*). The odorants included nine commonly recommended for neurological testing. For other sense modalities, the inherent confusability of stimuli seems to limit identification (*6*). The estimates for smell generally fall so low, however, as to suggest that factors other than inherent confusability limit identification. The four experiments reported here imply that sluggish acquisition and retrieval of odor names impede identification but that under the right circumstances confusability alone may set the upper limit. In the remarkably varied realm of odor quality, confusability poses only a minor limitation; when only this factor operates, persons can identify many odoriferous substances.

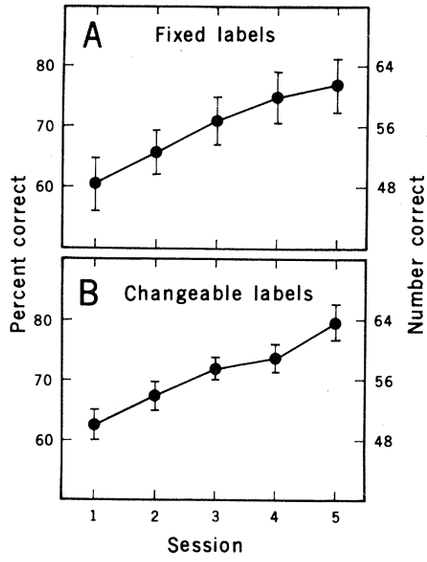


Fig. 1. (A) Percent correct and number correct when subjects sought to identify 80 substances with labels generated during previous inspection. Bars represent standard errors of the mean. (B) Similar to (A), except that the subjects had the option to change labels throughout testing.

In experiment 1, 12 women, blindfolded, sought to identify 80 commonly encountered, "ecologically valid" substances presented in irregular order from jars (*7*). Upon presentation of a substance, the subject first rated familiarity on a seven-point scale and then sought to name the substance. Average performance equaled 36 (range, 25 to 43). Moments after initial identification, the subject sought to "identify" the various substances again, but, on this occasion, sought to use only labels generated during the first exposure. Hence, in addition to asking how many substances a subject could identify veridically, the experiment looked at how consistently the subject could use her own labels, veridical or nonveridical. When incorrect on the second exposure, the subject received feedback regarding her previously generated label.


0036-8075/79/0202-0467$00.50/0  

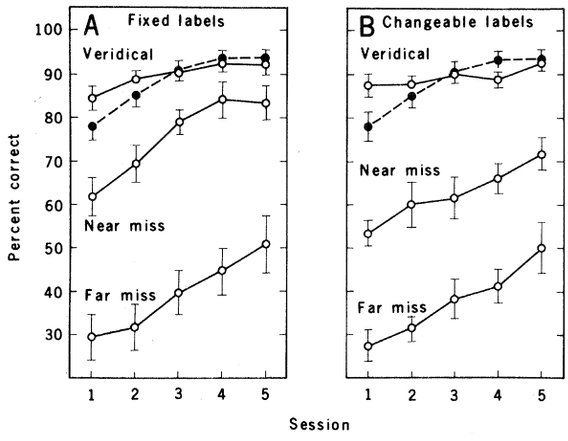


Fig. 2. (A) Open circles depict the data in Fig. 1A analyzed by quality of label (*20*). Closed circles depict how well subjects could identify all 80 substances when the experimenter disclosed the veridical labels. (B) Open circles depict the data in Fig. 1B analyzed by quality of label. Closed circles as in (A).

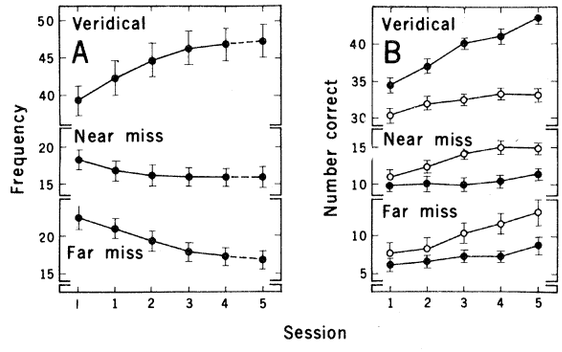


Fig. 3. (A) Effect of changing the label on the number in each qualitative category during the period ranging from inspection (*1*) to the final test. Only changes made through session 4 could have any bearing on identification, since switches of label in that session became the correct response for session 5. (B) Number of correct identifications, broken down by quality of label, for subjects permitted to change labels (closed circles) and subjects required to use fixed labels (open circles).

Each subject returned for four subsequent tests separated by about 2 to 3 days. On these occasions, she sought to identify the substances again with the labels generated previously and received feedback as before. This experiment shares features of a study by Engen and Pfaffmann, who also required subjects to identify odorants with previously generated labels (*8*). The present study differed, however, in its use of only frequently encountered substances with widely known names (such as chocolate, cinnamon, and peanut butter), rather than a combination of some such substances and some infrequently encountered ones without widely known names (such as pyridine, butanol, and acetone). In the earlier study, subjects could identify, after much practice and feedback, 17 substances.

When the subjects studied here sought to identify the substances shortly after initial inspection, performance equaled 60 percent (48.3 of 80) and climbed gradually to 77 percent (61.5 of 80) by the final session (Fig. 1). Presumably, such excellent ability to apply previously generated labels after only limited practice arose from the use of nameable, ecologically valid substances. The outcome actually surprises, however, in view of the many imprecise or incorrect labels generated during inspection. Although subjects who emitted a relatively high number of veridical labels performed better throughout testing ($r$ = .89), all subjects emitted many seemingly poor labels. This prompted the question, Did the quality of the label signal subsequent identifiability?

Two types of nonveridical labels seemed evident: (i) near misses (22 percent of initial labels), which comprised names of substances similar to and possibly confusable with test substances (for example, nutmeg for cloves, disinfectant for bleach) and (ii) far misses (33 percent), which generally included generic names (for example, fruit for orange) and specific but clearly incorrect names (cheese for machine oil), plus a few vague associations (Jimmy Carter for soy sauce) (*9*). Near misses, unlike far misses, had the earmarks of serviceable labels. A breakdown of performance by type of label substantiates this impression (Fig. 2).

Subjects rarely failed to identify substances to which they had assigned veridical labels. By comparison, most subjects infrequently identified substances which they had labeled with far misses, though a few subjects violated this rule consistently. These few, who were overall highly accurate, seemed able to use almost any label effectively no matter how worthless on the surface.

Substances judged high in familiarity during inspection generally received veridical labels, but some such familiar substances received near or far misses. The average rating for substances in the first category (veridical) equaled 6.0, whereas the ratings for those in the second and third categories equaled 5.1 and 3.5, respectively. Familiarity therefore influences degree of veridical identification, even among these substances chosen specifically for their common occurrence. But, familiarity hardly determines veridical identification; more than one-third of the ratings for substances labeled with far misses equaled the high values of 5, 6, or 7 (*10*). Often in the course of inspection, as in the everyday lives of most persons, a subject claimed to have the veridical name of some admittedly familiar substance on the tip of her tongue but could not emit the name (*11*).

Performance with near misses and far misses improved markedly over the five sessions. Does this mean that persons can readily learn to use an imperfect label more effectively? A second experiment offered evidence that an imperfect label may increase readily in usefulness only if accompanied by a covert verbal mediator of higher quality than the overt label. Another 12 women followed the same procedure as the previous group, but had the option to change labels. A subject could switch a label on any trial, but only after she had sought to identify the substance with its previous tag (and hence could be scored correct or incorrect) and only before she received feedback. The major question of interest: Would these subjects show progressive enhancement in the use of nonveridical labels or would they confine any progress to the use of better labels?

Subjects exercised the option to change labels 226 times. More than half the changes, 129, represented categorical improvements (for example, far miss converted to veridical), whereas only 30 represented categorical deteriorations. From the initial inspection to the final session, the number of veridical labels grew from 49 to 59 percent (Fig. 3). Hence, subjects frequently realized the exact identity spontaneously. Near misses fell from 23 to 20 percent, and far misses fell from 28 to 21 percent. Despite this net increase in the quality of the labels, overall performance was almost identical to that of subjects who used fixed labels (Fig. 1). Even a breakdown

of percentage correct by type of label would imply substantial agreement between the two groups (Fig. 2) (*12*). Nevertheless, a breakdown of number (rather than percentage) correct by label reveals that subjects permitted to switch labels made only trivial session-by-session increments in the number of substances identified with nonveridical labels (Fig. 3). Instead, they exhibited large increments in the number identified with veridical labels, as the pool of substances so labeled grew. Furthermore, the rate at which these subjects discarded nonveridical labels fell only slightly below the rate at which subjects in the other group seemingly "learned" to use such labels more effectively. These results imply that subjects required to use fixed labels commonly generated better labels covertly and used them as mediational links between the odor and the poorer fixed label. Indeed, subjects admitted this strategy freely. The strategy hardly surprises, but its apparent success, judged by the nearly identical overall performance of the fixed-label and changeable-label groups, points up the relative ease with which persons can form verbal-verbal links in contrast to olfactory-verbal links (*13*).

Imperfect identification during initial inspection could indicate in part merely poor learned associations between some odors and their names. It could also indicate more-or-less temporary inaccessibility of well-learned responses to odors. Evidence compatible with the second possibility (failure of retrieval) includes (i) spontaneous emission of better labels even when given corrective feedback regarding the label generated during inspection and (ii) an abrupt jump in performance with categorical improvements in labels. For instance, the probability of a correct response increased ninefold from the trial just before to the trial just after a switch from a far miss to a veridical label (*14*). The first two experiments revealed that such abrupt jumps do not characterize the mere strengthening of learned associations.

Insofar as inaccessibility of labels limits identification, then subjects should subsequently perform exceptionally well if, during inspection, they are merely prompted with (that is, informed of) the veridical labels. Under those circumstances, the percentage correct for all 80 substances should roughly match that registered for substances labeled veridically by the subjects themselves. Insofar as poorly learned associations between odors and names limit identification, subjects should show only a modest gain if informed of veridical labels during inspection and then given reinforcement (feedback) with the names during testing.

Another 12 women followed the procedure used previously, except that the experimenter named each substance during inspection and gave feedback with that name during testing. Performance on the first test after inspection equaled 78 percent (62.4 out of 80) and grew to 93.6 percent (74.9 out of 80) by session 5 (Fig. 2) (*15*). Such excellent performance supports the conclusion that blockage of retrieval limits performance when subjects must generate their own labels. On the assumption that subjects reinforced with veridical labels had long overcome any difficulties with retrieval by session 5, it seems reasonable to ask, Why did performance still fall below perfection? The answer seems to lie with inherent confusability, the most common limiting factor for the identification of stimuli outside the olfactory domain. In experiment 4, ten women discriminated among the substances by means of a reverse multiple-choice procedure. Each subject served twice, receiving no trial-by-trial feedback. On each of the 80 trials in a session, the experimenter first named a substance and then allowed the subject to smell three substances, including the correct one. Under these circumstances, which included appropriate randomization of all relevant factors, the subjects exhibited 94 percent discrimination in the first session and 95 percent discrimination in the second (both estimates corrected for guessing). Hence, at least a small fraction of the errors of identification seem attributable to failures of discrimination. In fact, the previous finding that once relieved of the burden of retrieval subjects reached a plateau at about 94 percent suggests that inherent confusability alone may have prevented perfect identification.

Hence, at least three factors impede odor identification: (i) sluggish formation of associations between odors and names, (ii) failure to retrieve the name in spite of a well-formed association, and (iii) inherent confusability of the stimuli. If tested with uncommon stimuli, such as laboratory chemicals, subjects will perform poorly because they possess neither associations between the odors and names nor the ability to develop them quickly (*16*). If tested with common stimuli, but required to name them initially with unaided recall, subjects will perform moderately well and will progress as better names come to mind. If tested with common stimuli and aided in recall, subjects will perform about as well as inherent confusability will allow. Knowledge of these factors can eliminate the notorious ambiguity in the norm for adequate performance in the standard neurological test of odor identification (*17*). When tested with 11 or 12 distinctive stimuli from the current set and aided in recall, persons with normal olfaction approximated or achieved perfect performance immediately and persons without olfaction scored zero (*18*).

For certain experts (perfumers, flavor chemists, food technologists, wine tasters), the pool of "common" odorants far exceeds that of the layperson. It therefore seems likely that these experts, who must frequently verbalize their olfactory experiences, would perform exceptionally well (*19*). Nevertheless, even the laypersons studied here generally felt that, if given aid in recall, they could have gone on to identify well over 100 substances (some said hundreds). This claim seems credible on the ground that the set of 80 stimuli included only a fraction of those odoriferous substances that most persons have smelled frequently and have slowly, but steadily, come to "know" with their noses.

WILLIAM S. CAIN
*John B. Pierce Foundation Laboratory*
and *Yale University,*
*New Haven, Connecticut 06519*

### References and Notes

1. D. Sumner, *Lancet* **1962-II**, 895 (1962).
2. F. N. Jones, in *Theories of Odor and Odor Measurement*, N. Tanyolaç, Ed. (Tanyolaç, Istanbul, 1968), pp. 133–141.
3. T. Engen and B. M. Ross, *J. Exp. Psychol.* **100**, 221 (1973).
4. J. A. Desor and G. K. Beauchamp, *Percept. Psychophys.* **16**, 551 (1974).
5. H. Lawless and T. Engen, *J. Exp. Psychol. Hum. Learn. Mem.* **3**, 52 (1977).
6. W. R. Garner, *Uncertainty and Structure as Psychological Concepts* (Wiley, New York, 1962), pp. 76–77.
7. The substances included (with examples) meats (bologna), fish (sardines), fruits (orange), spices (oregano), snacks (potato chips), condiments (barbecue sauce), beverages (grape drink), confections (caramel), household products (shoe polish), medical products (Band-Aids), raw materials (clay), personal products (baby powder), and other common items (leather, pencil shavings, crayons).
8. T. Engen and C. Pfaffman, *J. Exp. Psychol.* **59**, 214 (1960).
9. Any such classification poses the possibility of error. Most labels seemed to fall cleanly into one category or the other. Those few (perhaps 10 percent) that did not, led the experimenters to seek the advice of colleagues before the final decision.
10. The judgment of familiarity could reflect the frequency of an odor's occurrence in the subject's life, recency of occurrence, or a particularly memorable experience with an odor (for example, the smell of cloves, a topical anesthetic, during a painful dental procedure). The judgment could also reflect the subject's confidence that she could name the odor. This factor could have particular significance in these experiments, in which the subject had to name a substance almost immediately after she rendered a familiarity rating. In spite of this awkward and potentially embarrassing juxtaposition of naming and familiarity rating, subjects did give high familiarity ratings followed by vague or other-

wise incorrect labels. The subjects often remarked on these incongruities and made it apparent that a far miss only rarely carried with it a conviction that it was the veridical label.

11. Lawless and Engen (*5*) have termed this failure to retrieve the label for an odor the tip-of-the-nose phenomenon.
12. Correlation between number of veridical labels emitted during inspection and subsequent overall performance equaled .73.
13. There remains the question of why undiscarded nonveridical labels showed any serviceability whatsoever. The relatively high variability of the performance obtained with nonveridical labels (Fig. 2) reflects some of the reasons. Some nonveridical labels, even far misses, seemed to possess personal meaning (for example, Dad's bathroom) that endowed them with serviceability despite their surface imprecision. Other nonveridical labels apparently held neither personal meaning nor high surface precision and showed virtually no serviceability. In addition, even far misses often contained considerable generic information, as seen in the use of terms like spice for cinnamon, industrial chemical for turpentine, and so forth. Such generic terms led to erratic though hardly negligible identification. Finally, if a particular verbal-verbal link served well, the subject could merely choose to retain the nonveridical label [L. S. Prytulak, *Cognit. Psychol.* **2**, 1 (1971)]. Hence, nonveridical labels seemed to comprise a potpourri of personally meaningful, useless, and partially informative labels, and the high accompanying variability seems compatible with this diversity.
14. Abrupt jumps also accompanied switches from near misses to veridical labels (a factor of 2.2) and from far misses to near misses (a factor of 2.0).
15. Desor and Beauchamp (*4*) trained three subjects to identify almost 60 out of 64 odorants through use of a complex regimen of massed and distributed practice over many days using corrective feedback with veridical labels. The outcome of their experiment dispelled any doubt that laypersons could actually perform better than previously suspected, but left undetermined the reasons a training regimen succeeded.
16. Familiarity, though deliberately restricted here through the choice of common stimuli, played a role in all three of the present identification experiments. Average familiarity ratings correlated with subsequent identification in the following way: $r = .86$, .73, and .59 for experiments 1, 2, and 3, respectively ($P < .01$ throughout). Uncommon and hence unfamiliar stimuli would therefore seem to stand little chance of identification unless the subjects received long and arduous training. In fact, R. G. Davis [*J. Exp. Psychol. Hum. Learn. Mem.* **104**, 134 (1975)] found that subjects failed even to approach perfection after 20 trials when required to identify only four relatively unfamiliar odorants with numerals.
17. E. R. Bickerstaff, *Neurological Examination in Clinical Practice* (Blackwell, Oxford, 1968), p. 36.
18. W. S. Cain and J. Krause, *Neurol. Res.*, in press.
19. Jones (*2*), who permitted two perfumers to choose their own stimuli (perfume ingredients) and then tested identification with small sets of the chosen stimuli, estimated that such professionals could probably identify 100 to 200 odorants. Because the stimulus sets apparently excluded substances commonly encountered in the everyday lives of most laypersons, however, Jones's estimate may represent the increment that one type of professional experience can add to the relatively large number of common substances that laypersons can identify under the right circumstances.
20. The data depicted by the unfilled circles represent weighted averages. Hence, the breakdown of performance by the quality of label gave subjects who emitted more than average veridical labels a heavier weight in the calculation of the function for veridical labels and a lighter weight in the calculation of one or both of the other functions. As it turned out, this factor had virtually no net influence on the functions for veridical labels and near-miss labels. If all subjects were given equal weight in the final tally, irrespective of how many labels of each type they had emitted, these two functions would differ from those shown in the figure by less than 1 percent. On the other hand, the function for near-miss labels would rise by about 9 percent. The rise would reflect the increase in the relative contribution of those high-scoring subjects, noted in the text, who emitted fewer than average far misses but who could use virtually any label effectively. Nevertheless, the finding that the manner of computation would not change the functions for the veridical labels and near misses shows that the high performance obtained in these cases did not represent a statistical segregation of generally high-scoring subjects from low-scoring ones.
21. Supported by NIH grant ES-00592. This investigation began as a senior-year research project by R. Sax of Yale College. I thank him for his efforts in the early stages and H. G. Anderson III and R. J. Huey for technical assistance.

21 August 1978

## Regularity, Randomness, and Aggregation in Flowering Phenologies

Stiles (1) presented the results of a 4-year study on the flowering times of 11 hummingbird-pollinated plants in a Costa Rican rain forest. Stiles was primarily concerned with testing the hypothesis that "a system of compensating phenological responses of different species to unusual rainfall conditions may play a major role in maintaining an orderly, staggered sequence of flowering peaks among the hummingbird-pollinated plants." The basis of this hypothesis was the belief that natural selection should produce a regular sequence of flowering times, in order to minimize competition between plant species for pollinating hummingbirds or to minimize interspecific hybridization. Stiles concluded that "The phenological data . . . show that a regular sequence of flowering peaks was nearly always maintained . . . only during late November to early December was no hermit food plant ever at peak bloom. . . ."

The crux of Stiles' argument lies in demonstrating that the flowering times shown in his figure 1 are indeed regularly spaced. Stiles' conclusion that the pattern is regular within any one year is apparently based on a subjective examination of his data. One of us (B.J.R.) was faced with a similar situation in a study of flowering times in shrub communities. The flowering times observed in this study appeared regularly spaced, but unfortunately so did phenologies produced by assigning to each species a flowering time at random within the growing season. A subjective examination of the data was not sufficient to determine whether or not flowering times were, indeed, regularly, rather than randomly, spaced or even aggregated. Therefore, several statistical tests were developed to test the regularity hypothesis (*2*). One of these tests is applied below to Stiles' data. It shows that his sequences of flowering times in each of the 4 years are not regular as he concludes, but instead they tend to be aggregated in the drier parts of the year.

The null hypothesis is that the peak flowering date of each of the $k$ species is independently and randomly assigned a position along an axis representing the growing season from a rectangular (uniform) probability distribution. The length of the growing season is then normalized to one for computational simplicity (each peak flowering date is divided by the length of the growing season). These randomly assigned flowering peaks $x_1, x_2, \ldots x_k$ are then ordered from earliest to latest, designated as the order statistics of the sample $y_1, y_2, \ldots y_k$ where $y_1$ is the earliest flowering species and $y_k$ the last flowering species. The interval $y_{i+1} - y_i$ is then the distance in time between the peak flowering dates of any two adjacent flowering species. The null hypothesis is equivalent to the procedure of assigning to each of the $k$ species a peak flowering date at random from a table of random numbers, ordering the random numbers from first to last, and then normalizing everything to one. Given the null hypothesis, the statistical properties of $y_{i+1} - y_i$ can be derived (*2*). In particular, the mean of $y_{i+1} - y_i$ is $1/(k + 1)$, and the variance is $k/[(k + 1)^2(k + 2)]$.

Consider the sample statistic $P$.

$$P = \frac{\sum_{i=0}^{k} \{y_{i+1} - y_i - [1/(k+1)]\}^2}{k+1}$$

which is the sample variance of the distances between peak flowering dates between adjacent species, including the distance between the beginning of the growing season and the peak flowering date of the first species to flower and between the last peak flowering date and the end of the growing season. The expected value of $P$ under the null hypothesis of randomly assigned peak flowering dates is (*2*)

$$E(P) = \frac{k}{(k+1)^2(k+2)} \quad (1)$$

If peak flowering times tend to be regularly distributed through the growing season, the sample variance $P$ should be less than that expected from Eq. 1 (with 0.0 as a lower limit for perfect regularity); at the same time, if peak flowering dates are aggregated, the sample variance will exceed its expected value under the null hypothesis. The ratio $P/E(P)$ is, therefore, a measure of regularity or


0036-8075/79/0202-0470$00.50/0 