Journal of Chemical Ecology, Vol. 8, No. 7, 1982

Letter to the Editor

VERIFICATION VERSUS FALSIFICATION OF EXISTING THEORY
Analysis of Possible Chemical Communication in Crayfish

A disturbing feature in science is the frequent emphasis on verification of popular theories rather than on falsification of hypotheses. As Dayton and Oliver (1980) stressed recently "The verification of ideas may be the most treacherous trap in science, as counter-examples are over-looked, alternate hypotheses brushed aside, and existing paradigms manicured. The successful advance of science and the proper use of experimentation depend upon rigorous attempts to falsify hypotheses." While all disciplines of science suffer from this problem, the reliance of behavioral research on observational techniques requires that one exercise extreme caution in data interpretation. To avoid compromising the conclusions of field and laboratory studies, it is necessary to test rigorously alternative hypotheses and to rely on valid statistical techniques.

In his recent review of a 1981 paper by Itagaki and Thorp, Rose (1982) concluded that the earlier paper contained ". . . misconceptions concerning the nature of pheromones and intraspecific communication and misinterpretations of results within the paper." From our perspective the only potentially significant criticism concerned our general approach in evaluating experimental results. The opposite approach advocated at least de facto by Rose is illustrative of the problem mentioned previously. The specific criticisms by Rose and our opposite approaches to data interpretation are discussed below.

Although theoretically it takes only one case to reject a "properly framed" hypothesis, one must be sure that the results of a test are real (with regard to type I errors), exclusive of alternative hypotheses, and directly applicable to the overall question. The overall null hypothesis (H_0) in our study was that long-distance chemical communication of sexual identity, agonistic state, and stress condition does not occur among adult crayfish. To falsify the overall null hypothesis, it was necessary to show that (1) statistically significant results led to rejection of H_0, (2) these results were consistent with other data, and (3) the data were not equally well explained by alternative hypotheses. Two alternative hypotheses were that (1) the number of

0098-0331/82/0700-1073$03.00/0

comparisons declared statistically significant could have been explained by chance alone (type I error), and (2) similar results would have been obtained if we had employed novel stimuli (e.g., another taxa such as various fish species, etc.) rather than conspecifics. Because we did not conclude in our paper that "chemical detection" was equivalent to reception of species-specific pheromones, we did not proceed with contingency plans for employing novel stimuli. Unfortunately, proponents of the argument for pheromones in crayfish (e.g., Ameyaw-Akumfi, 1976) have failed to conduct the mandatory tests of this alternative hypothesis.

Significance levels are difficult to interpret in multitreatment experiments involving several simultaneous comparisons (Kirk, 1968, p. 82). Assigning a significance level is ambiguous because ". . . the conceptual unit can be the individual comparison, hypothesis, family of comparisons, or experiment. . . . The error rates become more divergent as the number of comparisons and hypotheses evaluated in an experiment are increased" (Kirk, 1968, pp. 82–83). In our study we tested 364 specific hypotheses based on 26 a priori comparisons (treatment groups, e.g., ABC, Table 2) for 14 "behaviors" (duration of 8 behaviors + average duration per trial for 3 behaviors + position in observation tank + right or left handedness = 14 total "behaviors"). Each of five basic experiments was replicated with 20 male and 20 female observation crayfish. We were aware that strict adherence to an alpha level of 0.05 for 364 hypotheses would produce a number of hypotheses falsely declared significant (type I error). The expected number of tests incorrectly declared significant would be approximately 18, which differed only slightly from the 30 statistically significant tests in our original Table 2 (for brevity the 334 nonsignificant results were deleted from the table). Of the 30 statistically significant hypotheses, only a few were a priori tests of communication of sexual identity; the remaining hypotheses were related to chemical detection (not communication) and communication of agonistic state and stress condition. Two multitreatment comparisons (ABC and CGH, Table 2) tested whether an observation tank male could distinguish between stimulus water from either controls (0 animals), 1 or 2 males, or 1 or 2 female crayfish. Only 1 of 28 hypotheses (14 behaviors each for the comparisons among ABC and CGH) was statistically significant compared to an expected 1.4 by chance alone. These results and others described in the 1981 paper led us to conclude that sexual recognition over long distances does not occur in crayfish through chemical communication.

The internal inconsistency of the results also lead us to reject the overall null hypothesis. For example, "Although females waved their chelae more often (DE) to solitary females than to solitary males, there were no differences in response to water conditioned by either two males or two females (IJ)" (Itagaki and Thorp, 1981). One would expect these internal inconsistencies if

the "statistically significant" results were random rather than real. Please note that we are not advocating deleting significant results from the publication; rather, we are recommending a conservative approach of not rejecting the overall null hypothesis (H_0) without strong and consistent results. In our case, we would have preferred to be able to falsify H_0 since that would have allowed us to pursue additional questions in this area, but the results did not justify rejecting H_0.

An alternative approach to the one advocated above is to accept all significant comparisons as real without regard to type I errors and to other problems mentioned previously. The effective result of this approach, whether intended or not, is to decrease the probability of falsifying H_0 and, thus, to increase the number of theories verified (just the opposite of the valid approach). If an investigator requires only one significant comparison to verify an hypothesis, it is almost impossible to falsify a theory. As an example of this nonconservative approach, Rose (1982) stated, "It seems more appropriate to argue that (1) *any statistically significant* [our italics] observable change is the expression of a real phenomenon given the constraints of the experimental regime, and (2) the behaviors showing such statistically significant differences are of *particular import*, [our italics] at least when compared to the others monitored . . ." We never stated or implied, as suggested by Rose, that all behaviors monitored in our study were of equal predictive value (in fact, meral spread was most representative of agonistic behavior, and it showed no significant responses). However, we disagree that it is appropriate, a posteriori, to emphasize only those comparisons or behaviors which are statistically significant. Ignoring non-significant results and emphasizing only significant tests biases one's conclusions.

In addition to critizing our data interpretations, Rose (1982) indicated that we had ". . . misconceptions concerning the nature of pheromones and intraspecific communication . . ." In drawing conclusions from experimental results, it is important to state the applicable "boundary conditions," such as the definition of the phenomenon investigated. Rose criticized our paper for not using the definition of communication that he had extracted from a 1951 edition of Webster's Dictionary. He also stated that we misquoted or misinterpreted the definition proposed by Wilson (1970); a direct comparison of our quote with Wilson's will show that this was not the case. Our use of the term "communication" follows the strict definition proposed by Burghardt (1970): "Communication is the phenomenon of one organism producing a signal that, when responded to by another organism, confers some advantage (or the statistical probability of it) to the signaler or his group." While Rose is free to interpret the significance of our conclusions in light of his colloquial definition, we reserve the right to use an established

scientific definition. It was not the purpose of our 1981 paper to analyze alternative definitions of communication; a thorough analyses has already been published (Burghardt, 1970). However, even if one accepts the nonscientific definition suggested by Rose, the perception by dogs of scent trails left by humans (Rose, 1982), would not qualify, in our opinion, as communication.

The remaining criticisms in Rose's paper concerned minor points in our original publication and are related, we believe, to his misinterpretations of our methods and/or assumptions. An explanation of all these points would be unnecessarily redundant here, but readers who are interested in further clarification should consult our 1981 paper. One point not stated in our earlier paper was that the position of observation tank crayfish in one of three zones was determined by the position of the eyes, not by the location of the entire body. Although we have never maintained that our flow-through apparatus was without limitations, we continue to believe that it has significant advantages over most static systems. The principal advantage relates to avoidance of toxic waste accumulation. Despite Rose's claim, we never stated or implied that pheromones could not be contained within excretory fluids (the scientific literature adequately demonstrates otherwise); rather, we believe that previous studies on adult crayfish communication (Ameyaw-Akumfi, 1976; Thorp and Ammerman, 1978) were not designed to distinguish between effects produced by toxic wastes or by possible pheromones in the excretory fluids of crayfish.

In conclusion, the nonconservative approach advocated by Rose (1982) ignores type I errors, is inadvertently biased in favor of statistically significant compared to nonsignificant results, and, as a result, emphasizes the verification of popular theories rather than the falsification of hypotheses. Rose's interpretation of what constitutes communication versus chemical detection is valid only if one accepts the colloquial definition from Webster's Dictionary over an established scientific definition. Although we found evidence of "chemical detection" by crayfish, we reaffirm the conclusion of our experimental study that ". . . chemical communication between adult crayfish does not occur or is not efficient at distances greater than the effective range for visual communication."

JAMES H. THORP
Savannah River Ecology Laboratory
P. O. Drawer E
Aiken, South Carolina

HARUHIKO ITAGAKI
Department of Zoology
Duke University
Durham, North Carolina

REFERENCES

AMEYAW-AKUMFI, C.E. 1976. Some aspects of breeding biology of crayfish. PhD dissertation, University of Michigan, Ann Arbor. 252 pp.

BURGHARDT, G.M. 1970. Defining "communication", pp. 5–18, *in* J.W. Johnston, Jr., D.G. Moulton, and A. Turk (eds.). Advances in Chemoreception, Vol. I, Communication by Chemical Signals. Appleton-Century-Crofts, New York.

DAYTON, P.K., and OLIVER, J.S. 1980. An evaluation of experimental analyses of population and community patterns in benthic marine environments, pp. 93–120 *in* K.R. Tenore and B.C. Coull (eds.). Marine Benthic Dynamics. University of South Carolina Press, Columbia.

ITAGAKI, H., and THORP, J.H. 1981. Laboratory experiments to determine if crayfish can communicate chemically in a flow-through system. *J. Chem. Ecol.* 7:115–126.

KIRK, R.E. 1968. Experimental Design: Procedures for the Behavioral Sciences. Brooks/Cole Publishing Co., Belmont, California.

ROSE, R.D. 1982. On the nature of chemical communication by crayfish in a laboratory controlled flow-through system. *J. Chem. Ecol.* 8:1065–1071.

THORP, J.H., and AMMERMAN, K.S. 1978. Chemical communication and agonism in the crayfish *Procambarus acutus acutus*. *Am. Midl. Nat.* 100:471–474.

WILSON, E.O. 1970. Chemical communication within animal species, pp. 133–155, *in* E. Sondheimer and J.B. Simeone (eds.). Chemical Ecology. Academic Press, New York.