Journal of Personality and Social Psychology
1979, Vol. 37, No. 3, 337–344

Ambient Temperature and the Occurrence of Collective Violence: A New Analysis

J. Merrill Carlsmith and Craig A. Anderson
Stanford University

Prevalent folklore suggests that riots tend to occur during periods of very hot weather. Baron and Ransberger examined 102 major riots in the United States between 1967 and 1971 and concluded that the frequency of collective violence and ambient temperature are curvilinearly related. The present article points out that the Baron and Ransberger analysis did not take account of the different number of days in different temperature ranges. The artifact is eliminated, and the probability of a riot, conditional upon temperature, is estimated. When this is done, the evidence strongly suggests that the conditional probability of a riot increases monotonically with temperature. Some general implications of such data analyses are discussed.

In a recent article in this journal, Baron and Ransberger (1978) presented an analysis of the relationship between the frequency of major riots and the ambient temperature occurring during the riots. To do this, they studied 102 major riots in the United States between 1967 and 1971. The hypothesis they wished to test, and for which they claimed confirmatory evidence, is the existence of a curvilinear relationship between the likelihood of a riot and the maximum ambient temperature at the time of the riot. This hypothesis contrasts with the prevalent folklore that riots tend to occur during periods of very hot weather. Specifically, Baron and Ransberger concluded that the likelihood of a riot increases with temperature up to the range of 81°–85° F and then decreases sharply with further increases in temperature. The evidence that they presented to support this relationship is a frequency distribution of the number of riots plotted against temperature. This frequency polygon does indeed peak in the interval 81°–85° F, falling off sharply on either side.

We contend that this relationship is an artifact of the particular way the data were examined and that an appropriate reanalysis suggests a monotonically increasing function relating the probability of riots and temperature. Basically, we argue that the Baron and Ransberger results stem from their having not taken account of base-rate differences in temperature. For example, if days in the 81°–85° F range are more common than days in the 91°–95° F range, there may well be more riots in the former range. There are, after all, many more opportunities for riots. But an appropriate analysis may well show that riots are *relatively* more common in the higher temperature range. To be sure, Baron and Ransberger did consider this possibility, but they rejected it. In our view, their rejection was premature; we consider their arguments and the weaknesses therein at greater length below.

The activities reported herein were supported in part by funds from Boys Town. However, the opinions expressed or the policies advocated do not necessarily reflect those of Boys Town. The research was carried out while the first author was a member of the Boys Town Center for the Study of Youth Development at Stanford University and while the second author was supported by a National Science Foundation Fellowship. We are indebted to Bradley Efron, Lee Ross, and Amos Tversky for their comments and suggestions on this work, and to Joy Davis for her assistance in the seemingly unending task of transcribing 58,000 temperatures.

Requests for reprints should be sent to J. Merrill Carlsmith, Department of Psychology, Stanford University, Stanford, California 94305.

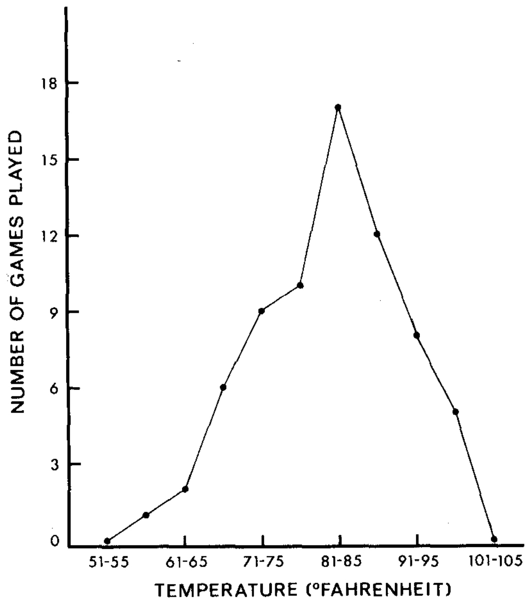

Figure 1. Frequency of baseball games played (New York Mets, 1977) as a function of ambient temperature.

Baseball, Temperature, and Base Rates

To see most clearly how this artifact may work, it is instructive to apply the same analysis used by Baron and Ransberger to a set of events that we know are not influenced by temperature. Figure 1 shows the same analysis used by Baron and Ransberger applied to the frequency of New York Mets baseball games played at home in 1977 (*The Sporting News,* April–October 1977). That is, we plot the frequency of Mets home games against the maximum ambient temperature in New York on the day of the game. (To be sure, baseball games occur primarily during the summer months, but then again, so do riots.) A brief study of Figure 1 shows a remarkable similarity to Figure 1 in the Baron and Ransberger (1978) article, and were we to follow their logic, we would have to conclude that "inspection of this figure lends support to the suggestion of a curvilinear relationship between ambient temperature and the incidence of [baseball games]" (p. 354). In view of the fact that baseball games are scheduled some months in advance, such a conclusion hardly seems warranted. Another explanation seems far more plausible. Both frequency polygons lead to erroneous conclusions, and for the same reason—the base-rate of different temperatures has not been taken into account. In our baseball example, it is not difficult to see that the frequency distribution of Mets home games is heavily influenced by the base-rate of daily maximum temperatures in New York during this period. Fewer games were played at temperatures of 91°–95° F, not because such temperatures lowered the probability of players choosing to play baseball, but because there were fewer such days. This simple example captures the essence of our critique of the Baron and Ransberger analysis, although the problem becomes a good deal more complex when we try to deal with the data they present.

The failure to consider base rates when assessing the probability of some event is hardly unique to this example. It is a problem of general concern in the analysis of data when we wish to calculate the probability of an event, conditional on the occurrence of some other event. It is also a problem that has begun to intrigue cognitive psychologists interested in subjective assessments of probability rather than formal statistical estimation. For example, in the context of a discussion of judgmental heuristics, Tversky and Kahneman (1974) comment:

> The reliance on heuristics and the prevalence of biases are not restricted to laymen. Experienced researchers are also prone to the same biases—when they think intuitively. For example, the tendency to predict the outcome that best represents the data, with insufficient regard for prior probability, has been observed in the intuitive judgments of individuals who have had extensive training in statistics. Although the statistically sophisticated avoid elementary errors, such as the gambler's fallacy, their intuitive judgments are liable to similar fallacies in more intricate and less transparent problems. (p. 1130)

Temperatures Before and After Riots (and Baseball Games)

The second major group of data presented by Baron and Ransberger is a plot of the maximum daily temperatures in the riot city during the 7 days prior to the occurrence of each riot and the 3 days following the riot. This plot shows gradually increasing tempera-

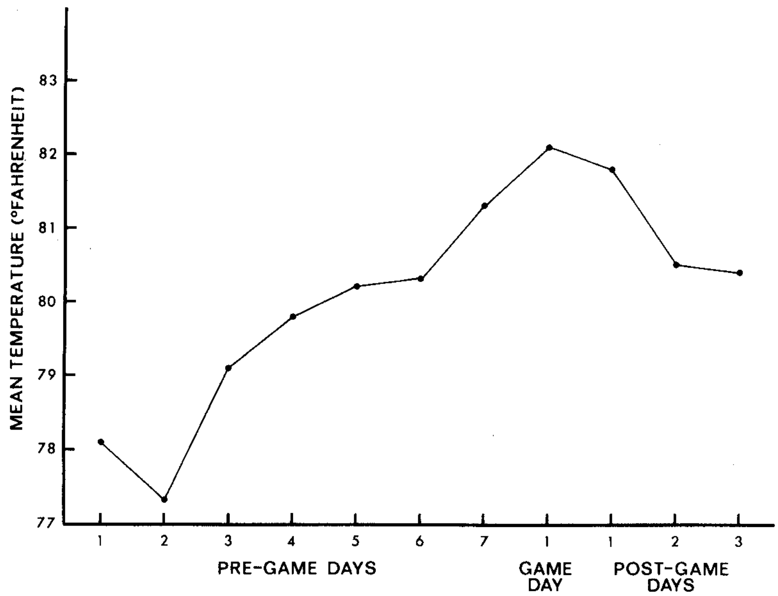

Figure 2. Mean maximum ambient temperature on the 7 days preceding the baseball game, on the game day, and on the 3 days following the game day.

tures up to the day of the riot, followed by decreasing temperatures during the 3 days following the riot. It should be noted that although this picture is consistent with their curvilinear hypothesis, it is also consistent with a wide variety of other possible functional relationships between temperature and the probability of riots. In particular, it is perfectly consistent with our hypothesis that there is a monotonically increasing relationship between temperature and the probability of a riot. On the other hand, the plot is equally consistent with a hypothesis that states that on occasional rainy (and cooler) days, riots (or baseball games) are unlikely to occur. Thus, the same picture could occur regardless of the true relationship between likelihood of a riot and temperature at the outbreak of a riot. Figure 2 shows a plot, again parallel to Baron and Ransberger's Figure 2, but done for New York Mets home games in 1977. Again we see a remarkable similarity between Baron and Ransberger's Figure 2 for riots and our Figure 2 for baseball games. It is our tentative hypothesis that our Figure 2 is mediated by a few rainy days, on which temperatures tend to be cooler and baseball games tend not to be played, but we are less certain of this artifact than we are of the artifact underlying Figure 1. In view of the fact that Baron and Ransberger's Figure 2 is consistent with a wide variety of possible functional relationships between temperature and the probability of riots, we will not pursue this issue further but return to the more fundamental question. If the postulated curvilinear relationship between temperature and the probability of riots is artifactual, what is the nature of the true relationship?

Conditional and Unconditional Probabilities

It is easier to point to the dangerous artifact underlying Figure 1 than to see a perfect solution to it. We present two alternative analyses below. Neither is immune to criticism, although both remove the obvious artifact. The consistency that emerges from these two very different analyses leads us to some confidence in the conclusions they imply, although we would emphasize at the outset our qualms about drawing any firm conclusions from this type of correlational analysis.

To see how to remove the effect of base rates, it is instructive to formalize our discussion slightly. The quantity we wish to estimate is the *conditional probability* of a riot, given a particular temperature. Since we have only 102 data points with which to work (or 86 if we follow Baron and Ransberger in analyzing separately the 16 riots that occurred on the date or anniversary of the Martin Luther King assassination), we can only hope to estimate this conditional probability of a riot given that the temperature is in a particular interval. Let E_i correspond to the event that the temperature is in the ith interval. Then $P\{R|E_i\}$is the conditional probability of a riot given a temperature in the ith interval. We follow Baron and Ransberger in using 5° intervals. From a familiar relationship in elementary probability theory, we have the following equation:

$$P(R|E_i) = \frac{P(R \,\&\, E_i)}{P(E_i)}.$$

That is, the conditional probability of a riot, given that the temperature is in a particular interval, is given by the joint probability of a riot and a temperature in a particular interval, divided by the probability that the temperature is in that interval.[1] Examination of the equation makes it clear that Baron and Ransberger essentially estimated the joint probability without correcting for the marginal temperature distribution.

Our problem, then, is to estimate that marginal distribution. It is not such an easy problem as it might appear, since the universe from which the particular riot temperatures are to be viewed as a sample is not well defined. We might attempt to conceptualize it as the distribution of all temperatures in the United States in the 5-year period in question. But a moment's thought shows the vagueness of that conceptualization. Should that distribution be weighted by the population density in each geographical location? Does it include the temperatures at Death Valley, where there are too few people to stage a convincing riot? Alternatively, we might try the distribution of temperatures in cities larger than, say, 100,000 people, again over the 5-year period in question. But some riots occurred in much smaller cities. Furthermore, no riots occurred in Alaska. Should we then include Alaska in our universe? Our solution to this problem was to define the universe of temperatures as all temperatures occurring in the 79 cities in which there were riots over the 5-year period in question.[2]

This definition generated a 79 × 1,826 matrix of daily temperatures. In order to estimate the distribution of temperatures in this matrix, we randomly sampled 2 of the 5 years for each city and found the daily maximum ambient temperature for each of the 57,705 days so defined. Temperatures were obtained from the *Climatological Data* reports of the U.S. Department of Commerce (Environmental Data Service, 1967–1971). Where no reporting station existed in the city (as was true in 6 of the smaller towns), we used the nearest station. This method, then, yielded an estimate of the probability distribution of maximum daily temperature over all riot cities over the 5-year period in question. That is, for each 5° F temperature interval, we had a count of the number of days in which the maximum temperature was in that interval. Dividing that count by 57,705 (the total number of days sampled) yielded an estimate of the probability of a day in

---

[1] An urn model may serve to clarify this point. Suppose we imagine an urn with a large number of marbles, each marked with a temperature. Most of the marbles are white, but a few are blue. The blue marble corresponds to a riot. To estimate the marginal distribution of temperature, we draw a sample of marbles and plot the temperatures. To estimate the probability of a riot, we draw a sample of marbles and count the number of blue marbles. To estimate the joint probability of riots and temperature, we draw a sample of marbles and count the number of blue ones in each temperature range. To estimate the conditional probability of a riot given a particular temperature, we draw a large sample, count the number of blue marbles in a particular temperature range, and divide that count by the total number of marbles in that temperature range. This final number is interpreted as follows: If we are told that we have a marble in a particular temperature range, the conditional probability tells us the number of chances that it is blue. So we wish to know for our data, Given that the temperature is between 81° and 85° F, what is the conditional probability of a riot? Having estimated this conditional probability for each interval, we then ask how these conditional probabilities vary with temperature.

[2] We are indebted to Robert A. Baron for providing us with the list of cities and dates of the 102 riots used in their study.

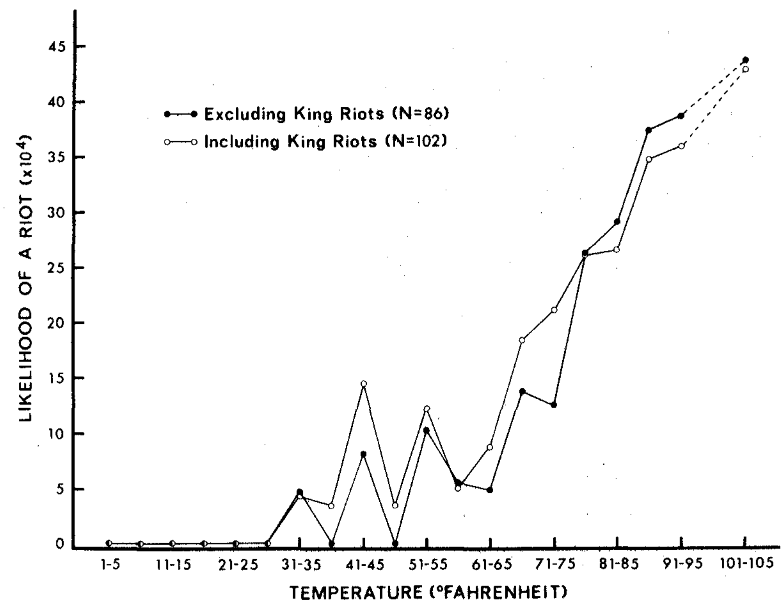

Figure 3. Conditional probability (likelihood) of a riot as a function of ambient temperature.

which the maximum temperature was in that interval. To estimate the conditional probability of a riot, given that the temperature was in a particular interval, we calculated the number of occurrences of a riot with a temperature in that range and divided by the number of days with a temperature in that range. The resulting function is shown in Figure 3.[3]

It is apparent that Figure 3 provides no evidence for a curvilinear relationship. Instead, there is a continuously increasing likelihood of riots as the temperature continues to rise, at least up through the temperature range of 91°–95° F. Contrary to the assertion of Baron and Ransberger, riots do not seem to be most likely at temperatures between 81° and 85° F; rather they become more and more likely with increasing temperature. This different function, of course, is a consequence of the fact that although there are fewer riots on days when the temperature is, say, 91°–95° F than on days when the temperature is 81°–85° F, there are many fewer days in the higher temperature range. The smaller number of riots at extremely high temperatures appears to be the result of many fewer opportunities for riots to occur; the conditional probability of a riot is larger at the higher temperature.

Three brief methodological notes are in order about Figure 3. First, there is a total of only three riots in the highest three temperature intervals, making the probability estimates extremely unstable. Consequently, we have averaged the three points, and this average is connected to the remainder of the function by a dashed line. Second, in defining the maximum daily temperature associated with each riot, we used the temperature on the day the riot began. This method contrasts with that of Baron and Ransberger, who took the average of the daily temperatures over the duration of the riot, for those riots that lasted more than 1 day. It seemed to us that if one wants to consider temperature as a causative factor in the outbreak of riots, it is more sensible to measure the temperature at the time of the outbreak rather than to include temperatures over subsequent days. Clearly, the temperature on

[3] We have followed Baron and Ransberger in estimating each function twice—once including the Martin Luther King-related riots and once excluding them.

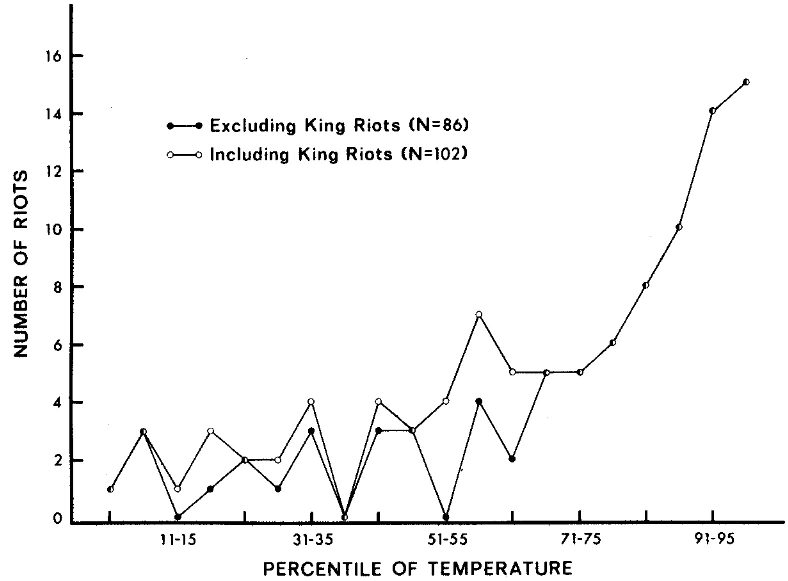

Figure 4. Frequency of collective violence (riots) as a function of the relative ambient temperature (expressed in percentiles).

July 12 cannot be more influential in determining whether a riot begins on July 10 than is the temperature on July 10. If we were to follow the Baron and Ransberger definition, there would be some minor changes in our Figure 3. Since we know that temperatures following the onset of a riot tend to be less extreme than temperatures on the first day of the riot, it is not surprising to find that extreme temperatures are slightly less common if we average over the days of the riot and that the proportion of riots at the modal temperature increases slightly. Even were we to use this averaging, the function still fails to show the precipitous drop that made Figure 1 so compelling, and the riot probability still reaches its maximum in the highest temperature range. Third, we present no inferential statistics in conjunction with this figure. It is our view that such statistics are, at best, irrelevant to these data and, at worst, seriously misleading. In their analysis of Figure 1, Baron and Ransberger present chi-square statistics. As we have already seen, such a calculation rests on an assumption that all temperature intervals are equally likely—an assumption that is demonstrably false. Furthermore, the riots show strong temporal and geographical dependencies (for example, five riots occurred in different cities in Michigan in a 3-day interval), which make assumptions of independence untenable.

An Alternative Analysis

Although Figure 3 casts severe doubts on the curvilinearity hypothesis, it is not totally convincing by itself. We have already alluded to the somewhat arbitrary definition of the universe of temperature days. Furthermore, the base-rate information used in the calculation of Figure 3 can be overly influenced by an extreme temperature distribution in one or two cities (although the large number of data points makes this somewhat implausible). Still, we are adding across rather dissimilar temperature distributions. A quite different analysis, which avoids these particular difficulties, involves looking at each riot temperature relative to all temperatures in the riot city. Thus, we estimate the cumulative distribution function (cdf) for all temperatures in City K and then look at the temperature on the day of the riot in that city relative to the cdf of all temperatures in that city. This procedure converts each riot temperature to a percentile relative to all temperatures in the riot city. Again, we use the random sample of 2 of the 5 years in each city over the period 1967–1971 to estimate the cdf of all temperatures in that city, estimating each cdf by 730 (nonindependent) points.

Figure 4 shows the number of riots occurring at each percentile interval. By shifting to percentiles, we have eliminated the base-rate problem that plagued us in Figure 1. When we plot the frequency of riots occurring at different percentiles, the null hypothesis clearly predicts a uniform distribution. Any effect of the different likelihoods of different temperatures has been removed by expressing each riot temperature relative to all temperatures in that city. Figure 4 hardly suggests a uniform distribution, nor is there any evidence of curvilinearity. What is suggested is a monotonically increasing function. Thus, once again we conclude that the likelihood of a riot in a given city increases as the maximum ambient daily temperature in that city increases.

It should be noted that Figure 4 examines the covariation of riots and the relative temperature (relative to all temperatures in that particular city) rather than the riots' covariation with the absolute temperature. The relationship between relative and absolute temperature is strong enough (although by no means perfect) that we see no hope of using these data to answer the fascinating question of whether absolute temperature or relative temperature (relative to some adaptation level) is more important in predicting the likelihood of a riot. Rather, we see the convergence of these two functions, based on very different methods of analysis, as lending support to the general proposition that, at least for these particular riots, the probability of a riot increases monotonically with increasing maximum ambient temperatures in potential riot cities.

Attempts to Discount the Base-Rate Artifact

In view of the dramatic differences in conclusions that we draw after taking account of different temperature base rates, it is worth considering why Baron and Ransberger rejected this artifact as an explanation of their results and why we feel that the rejection was in error. They present several arguments against the plausibility of the artifactual effect, but each is particularistic and fails to take into account the overall impact of the base rate. Analogous to our use of Mets baseball games, for example, they looked at celebrations following victories in particular major athletic contests (bowl games, Stanley Cup play-offs, National Basketball Association play-offs, and the World Series) and found that temperatures associated with such events do not peak at 81°–85° F. Unfortunately, none of the events they chose to study occur in the summer months. Thus, they failed to observe that over the course of the year in the riot cities, temperatures in the 81°–85° F range are indeed more frequent than those in any other interval. The only evidence they present relevant to the overall distribution comes from their second counterargument against it. They select 11 cities and 2 months (July and August) and show that temperatures in the 81°–85° F range are not uniformly most frequent in those cities in those months. But different choices of cities, months, and temperature intervals lead to different conclusions, none of which describe the overall temperature distribution. As we have shown above, temperatures in the 81°–85° F range are in fact the most common temperatures in these cities. Baron and Ransberger's other arguments have this same particularistic quality, focusing primarily on temperatures in particular ranges in riot cities on the same dates in riot years versus nonriot years, and we do not consider them in detail here.

A Final Note

We close with a final set of cautionary remarks. We feel quite confident that these data do not provide support for the hypothesis of a curvilinear relationship between temperature and the probability of a riot. We feel reasonably confident that for these particular riots, there is good evidence for a monotonically increasing relationship between temperature and the probability of a riot.[4] However, facile generalizations from

[4] These remarks about the monotonic character of the relationship between temperature and the likeli-

these data make us very nervous. The riots are certainly not independent of one another; the dependencies cannot easily be described. The temperatures are not independent; again the dependencies are complex. Thus, from the point of view of inferential statistics, the number of true independent data points is unknown and may be small. Not only are the data fraught with all of the ambiguities of any correlational study, the data analyses are also subject to subtle, difficult, and complex effects. Even if the present data analysis seems satisfactory, there are numerous alternative explanations of the relationship. A clear understanding of the psychological effects of temperature, and particularly the effects of temperature on aggression, seems much more likely to emerge from experimental work like that of Baron (1972), Baron and Bell (1976), or Baron and Lawton (1972).

hood of riots are restricted to the normal range of temperatures. Clearly, at some point the relationship must become curvilinear. We seriously doubt that riots are likely to occur when the temperature is 120° F (although we have no data one way or the other).

References

Baron, R. A. Aggression as a function of ambient temperatures and prior anger arousal. *Journal of Personality and Social Psychology,* 1972, *21,* 183–189.

Baron, R. A., & Bell, P. A. Aggression and heat: The influence of ambient temperature, negative affect, and a cooling drink on physical aggression. *Journal of Personality and Social Psychology,* 1976, *33,* 245–255.

Baron, R. A., & Lawton, S. F. Environmental influences on aggression: The facilitation of modeling effects by high ambient temperatures. *Psychonomic Science,* 1972, *26,* 80–82.

Baron, R. A., & Ransberger, V. M. Ambient temperature and the occurrence of collective violence: The "Long, hot summer" revisited. *Journal of Personality and Social Psychology,* 1978, *36,* 351–360.

Environmental Data Service. *Climatological data.* Washington, D.C.: U.S. Department of Commerce, 1967–1971.

The Sporting News. April, 1977, through October, 1977.

Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science,* 1974, *185,* 1124–1131.

Received June 29, 1978 ■