Notes

1. A zero-sum game is one in which, regardless of the outcome of the game, the winnings of one player are exactly balanced by the losses of the other, or others, i.e., the algebraic sum of the "payoffs" to each player is always zero. If this sum is always constant, not necessarily zero, regardless of the outcome of the game, the game is called "constant sum." From the point of view of game theory, constant-sum games are strategically equivalent to zero-sum games. We shall refer to all constant-sum games as zero-sum games.

2. Nonzero-sum games are all those which are not constant-sum. By definition, then, some outcomes of nonzero-sum games are jointly better for *both* players than other outcomes (assuming that the payoffs can be added). Even if the payoffs cannot be added, it is possible for both (or all) players to prefer some outcomes to others, in contrast to two-person zero-sum games, in which the preferences of the two players must always be opposite.

3. The simplest self-negating proposition is "Every statement within these quotation marks is false." If one assumes the proposition true, it turns out to be false; if one assumes it false, it turns out to be true.

4. This is in consequence of the so-called Law of Large Numbers, which states that if a random variable assumes different values with fixed probabilities, then the average of the values it has assumed over a long sequence of realizations will be very nearly equal the sum of the products of the several values, multiplied by the respective probabilities (the expected value). In the case under consideration, the expected value is $1/2(1) + 1/2(-1) = 0$.

5. Smallest and largest are meant here in the algebraic sense. Of two negative numbers, the numerically larger one is the "smaller."

6. We shall be designating outcomes by pairs of strategy choices which determine them. Thus the outcome (C_1C_2) is the result of player 1's choice of C_1 and player 2's choice of C_2. When

there is no danger of ambiguity, we shall sometimes omit the subscripts. Thus we shall sometimes write (CD) for (C_1D_2), etc.

7. When a game is in normal form, one strategy is said to dominate another for a player, if all the outcomes in the former strategy are at least as good for the player in question and sometimes better than the corresponding outcomes in the latter. That is, no matter what strategy the other player chooses, the dominating strategy yields a larger payoff than the dominated strategy.

8. Note that strategy 2 seems rather reasonable. It starts out by "trying" cooperation. The second move is contingent on the result of the first. If cooperation has worked, it is continued, but not otherwise. These, however, are commonsense arguments. They do not withstand a more thorough strategic analysis according to which it is better to play D in the second move regardless of what happened on the first.

9. Lewis F. Richardson investigated the properties of mathematical models proposed to represent arms races. The most interesting property of these models is their instability when the parameters assume values in a certain range. An arms race represented by a model in the unstable range must either keep going at an accelerated pace (which Richardson interpreted as a drive toward war) or must go into reverse, i.e., toward disarmament. We shall refer to this as the Richardson Effect (cf. Chapter 8).

10. Strictly speaking, when a game is represented in normal form, the payoffs designate the utilities of the associated outcomes to the players. Utilities are so defined that preferences of players with respect to risky outcomes (i.e., preferences among "lottery tickets," each representing an assignment of probabilities to each of the possible outcomes) is determined by maximizing expected utility. Utilities are supposed to be given on an interval scale. That is, if all the utilities u_i of either player are changed from u_i to $Au_i + B$ $(A > 0)$, the resulting game is supposed to be identical with the original one. In our experiments the payoffs were in money. We have not, of course, undertaken the very difficult task of determining a utility scale of money for each subject. We have simply assumed that the money payoffs were utilities. In most cases, this is the best one can do. In our situation, the determination of utility scales was not necessary, because we were not testing any game-theoretical result which relates specifically to utilities.

11. When more than two persons are involved, it is sometimes necessary to assume that payoffs are transferable and additive. In that case, it is assumed that payoffs are identified with utilities in the absolute sense.

12. To see this, consider a simpler case, involving only three variables. Let $F(x,y,z) = F(ax + b, ay + b, az + b)$; in other words, let a function of three variables retain its value if all three variables are subjected to a linear transformation. Let $p = (x - y)/(y - z)$, $q = y - z$. Then, given values of p, q, and z, the values of x, y, and z are determined since $y = q + z$, $x = pq + q + z$. But then we can write $F(x,y,z) = G(p,q,z) = G(p,aq,az + b)$ where G is some other function of three variables. Let $a = 1$. Then $G(p,q,z + b) = G(p,q,z)$, which shows that G is independent of z, because b is arbitrary. We can therefore write $G(p,q,z) = H(p,q)$. But then $H(p,q) = H(p,aq)$ and since a is arbitrary, we must have H independent of q. This leaves $p = (x - y)(y - z)$ as a single variable in which $F(x,y,z)$ can be expressed. In other words, if a linear transformation on the variables leaves a function of those variables invariant, the function can be expressed as a function of ratio differences. In the case of three variables, one such ratio suffices. In the case of four variables, two are needed.

13. This is in consequence of the fact that the product moment correlation coefficient is invariant with respect to any linear transformation.

14. The equations of a Markov chain indicate how the probabilities of responses depend on what these probabilities were in the preceding play. Thus our two simpletons will play CC on a given response if either (1) they played CC on the preceding response and neither defected from it, or (2) they played DD on the preceding response. Denoting by $(CC)'$ the probability that CC will occur on the next response, we have the equation

$$(CC)' = x_1x_2(CC) + DD.$$

Similarly,

$$(CD)' = x_1(1 - x_2)CC,$$
$$(DC)' = x_2(1 - x_1)CC,$$
$$(DD)' = (1 - x_1)(1 - x_2)CC + CD + DC.$$

If a steady state is reached, we must have $(CC)' = (CC)$, $(CD)' = (CD)$, $(DC)' = (DC)$, $(DD)' = (DD)$. It is shown in the theory of stochastic processes that in this case a steady state will eventually be approached arbitrarily closely. Setting $(CC)' = (CC)$, etc., we get four equations in four unknowns, namely the asymptotic probabilities of the four states (actually three equations in three unknowns, since the four probabilities must add up to 1). These equations have a unique solution given by Equations (20)–(23) of the text.

15. A partial derivative of a function of several variables with respect to one of these variables is the rate of change of the function as the variable in question varies while the others remain constant. Clearly this rate of change depends in general on the (fixed) values assigned to the others.

16. Cf. Bush and Mosteller, 1955, p. 109.

17. For example, if we are asked to pass a straight line through any two given points, we can always do it: it takes two parameters to determine a straight line given by the equation $y = mx + b$. If we have three points not all on the same straight line, we can always draw a parabola through them, for a parabola is determined by three parameters, etc.

18. A differential equation is one which relates the rates of change of variables to the variables themselves. For example, a population reproducing at a constant rate per individual is described by the differential equation $dp/dt = kp$. The equation says that the rate of growth of the whole population is proportional to the size of the population.

    A solution of a differential equation is obtained when the derivatives (rates of change) have been eliminated and only relations among the variables themselves remain. For example, the solution of the above differential equation is $p = p_0e^{kt}$, where p_0 is the size of the population at time 0 ($t = 0$), $e = 2.718$ approximately, and k depends on the rate of increase per individual. In physical science the most important solutions of differential equations are those which describe how certain variables change with time, i.e., describe processes. In our case, too, we are interested in the way our variables (e.g., *C*, *CC*, *CD*, etc.) change in the course of an experimental session.

19. Recall, for example, that if we assume that the payoffs are given only on an interval scale, i.e., if the zero point and the unit of payoffs must remain arbitrary (as is usually assumed in the theory of the two-person game), we are not free to choose any function of the payoffs as a "cooperative index" of a game. We are confined to functions in which only the ratios of differences of the payoffs appear as independent variables.

20. This is what we did in the case of the two simpletons (cf. Chapter 4 and Note 14).

21. An expression is indeterminate for some value(s) of its variable(s) if substituting the value(s) results in an expression like $0/0$ or ∞/∞. Nevertheless, the limit of such an expression can be often determined as the critical value(s) is (are) approached. For example, the expression $(x^2 - 1)/(x - 1)$ becomes $0/0$

when $x = 1$. However, as x becomes arbitrarily close to 1, the value of the expression becomes arbitrarily close to 2. Limits of this sort can be evaluated by calculating the ratio of the derivatives (rates of change) of the numerator and the denominator with respect to the variable in question (Hôpital's Rule).

22. $F(1,1)$ designates the value taken by the expression $F(x_1,x_2)$ when $x_1 = 1$, $x_2 = 1$.

23. Recall that the expected payoff is the sum of all possible payoffs multiplied by their respective probabilities of occurrence, which are the outcome probabilities, CC, CD, DC, and DD.

24. Expressions (75) and (76) vanish if and only if their numerators vanish.

25. If the expression is not to exceed 1, the numerator must not be greater than the denominator.

26. Dividing through by R^2, rewrite (81) as $3Q^2 - 8Q - 3$, where $Q = T/R$. There is no upper limit to Q, but 1 is the lower limit, since $T > R$. When Q is large, $3Q^2 - 8Q - 3$ must be positive, and so must expression (81). Q vanishes when $Q = 3$ and remains negative as Q decreases to its lower limit, 1.

27. An "operator" can be viewed as a set of directions which indicate how one quantity is to be transformed into another. In this case the directions are: "To change $p(t)$ into $p(t + 1)$, multiply $p(t)$ by α, then add $(1 - \alpha)\lambda$." In a given context the general form of the directions is the same, but the quantities α and λ may change. They are therefore the parameters of the operator.

28. The order of a differential equation refers to the highest order derivative occurring in it. Thus second order differential equations involve not only rates of change of variables but also rates of change of rates of change. In mechanics these are usually accelerations of masses. In electrodynamics the second derivative of the electric charge is related to the electromotive force.

29. Strictly speaking, we are dealing here not with continuous rates of change dC/dt but with rather finite discrete increments. However, if the increments of time are also considered small compared to the total time elapsed, the ratios of the increments can be considered as derivatives. The advantage of differential equation models is that they can be handled by well-known methods.

30. For that is the pair of values of C_1 and C_2 where their rates of change vanish, hence the values are stationary—a definition of equilibrium.

31. Proof of (1). In this case, as we have seen, $F(C)$ has no real roots. But $F(0) > 0$. Therefore $F(C)$ is always positive, and so is dC/dt. Therefore, C will be always increasing.

Proof of (2). In this case $F(C)$ has two real roots, namely

$$r_1 = \frac{(\beta + \gamma + 2\delta) + \sqrt{(\beta + \gamma + 2\delta)^2 - 4(\alpha + \beta + \gamma + \delta)\delta}}{2(\alpha + \beta + \gamma + \delta)},$$

$$r_2 = \frac{(\beta + \gamma + 2\delta) - \sqrt{(\beta + \alpha + 2\delta)^2 - 4(\alpha + \beta + \gamma + \delta)\delta}}{2(\alpha + \beta + \gamma + \delta)}.$$

If $(\alpha + \beta + \gamma + \delta) > 0$, it can be easily shown by comparing the numerator with the denominator and the two terms of the numerator that both roots lie between zero and one. Hence this is the same case as is represented by Equations (99) and (100) of the text.

Proof of (3). In this case two real roots of $F(C)$ are guaranteed, because the expression under the radical is positive. We shall show, however, that only the larger root lies between zero and one. Suppose first that $(\alpha + \beta + \gamma + \delta) > 0$ so that the denominator of r_1 and r_2 is positive. Then r_1 is the larger root. We shall show that $r_1 \leq 1$ if we show that

$$(\beta + \alpha + 2\delta) + \sqrt{(\beta + \alpha + 2\delta)^2 - 4(\alpha + \beta + \gamma + \delta)\delta} \leq 2(\alpha + \beta + \gamma + \delta).$$

This is true if

$$(\beta + \gamma + 2\delta)^2 - 4(\alpha + \beta + \gamma + \delta)\delta \leq (2\alpha + \beta + \gamma)^2,$$

which is true if

$$(\beta + \gamma)^2 - 4\alpha\delta \leq (\beta + \gamma)^2 + 4\alpha(\beta + \gamma) + 4\alpha^2,$$

which is true if

$$0 \leq 4\alpha(\alpha + \beta + \gamma + \delta).$$

But this is true, because we have assumed that

$$(\alpha + \beta + \gamma + \delta) > 0, \alpha > 0.$$

It is clear that $r_1 > 0$, because both its numerator and its denominator are positive.

Next suppose that $(\alpha + \beta + \gamma + \delta) < 0$ so that the denominator of r_1 and r_2 is negative. Now r_2 is the larger root. By a similar argument we prove that now r_2 lies between zero and one, while r_1 is negative.

In either case, then, the larger root lies between zero and one and is the only such root of $F(C)$. But this is the root associated with the unstable equilibrium. Thus (3) is proved.

32. The situation is analogous to that in classical thermodynamics where a physical or chemical system is assumed to go through a reversible process, i.e., a process so slow that the equilibrium is always reestablished following each minute change. While such a situation is conceivable in the case of certain chemical reactions, for example, where the changes are slow relative to the time it takes for the equilibrium to be reestablished, it is quite unlikely that analogous processes can be found in psychological interactions.

33. Note that the high frequency of C observed in Game II (Pure Matrix Condition) is attributed to the large numerical magnitude of $P(= -9)$. But in our model the punishment at DD has little effect, because escape from DD is immediate ($w = 1$).

34. To see this, consider a population of runs with a constant mortality. Then the number of runs "dying" at a given moment is proportional to the number surviving. This is expressed by the differential equation

$$-\frac{dP}{dt} = mP,$$

where P is the fraction of the population surviving. The solution to this equation is

$$P(t) = P_0e^{-mt},$$

where P_0 is the fraction surviving at $t = 0$. But this fraction is the entire population at $t = 0$. Hence $P_0 = 1$, and we obtain Equation (120) of the text.

35. Taking negative logarithms of both sides of

$$P(t) = e^{-f(t)},$$

we get

$$-\log_e P(t) = f(t).$$

If $f(t)$ is linear in t, the plot of $-\log_e P(t)$ against t will be a straight line. If $f''(t) > 0$, the plot will be concave upward. If $f''(T) < 0$, the plot will be concave downward.

36. We were at first intrigued by the good fit given by the formula, because many apparently self-enhancing processes are also well described by it, e.g., distribution of durations of wars and strikes (Horvath and Foster, 1963).

37. The Laplace transform of a function $f(t)$ is by definition $\int_0^\infty f(t)e^{-st}\,dt$. The variable of the integration t integrates out, leaving the parameter s as the argument of the Laplace trans-

form. In our case, μ plays the part of t while t plays the part of s.

38. Examination of the seven games separately reveals that women cooperate less in every one of the seven games.

39. In mixed pairs, M must be computed as follows: $(1 - y_1)(1 - z_2)/y_1z_2$ is associated with player 1. Thus if the man is player 1, this M is the ratio of failures to successes of men's martyr runs. If the woman is player 1, it is the corresponding ratio related to women's martyr runs.

40. Pilot studies performed in preparation of the next phases indicate that cooperation in response to the stooge's one hundred percent C in Game III is not minimal but clusters at the extremes. Some subjects cooperate almost completely; others defect almost completely. The latter, however, seem to be in a majority.

41. A tit-for-tat strategy implies $\rho_1 = 1$ (cf. pp. 61–62).