BLACKJACK SYSTEMS*

by

Edward O. Thorp
Mathematics Department
University of California
Irvine, California

ABSTRACT

A simple formula is presented for approximately evaluating the betting effectiveness of an arbitrary point count system. A simulation procedure is proposed and described and it is shown that this procedure determines with statistical accuracy the maximum possible theoretical advantage from counting all cards. This is the advantage from perfect (computer) play, assuming random shuffling. The procedure can be used to separately determine the portions of the maximum theoretical advantage from (1) using the basic strategy and (2) by incorporating strategy variations.

Two new classes of possibly favorable and practical blackjack systems are discussed. One is based on non-random shuffling. The other is based on non-random "stopping".

*Paper presented at the Second Annual Conference on Gambling. South Lake Tahoe, Nevada, June 15-18, 1975.

BLACKJACK*

1. Introduction. Practical winning strategies for the casino blackjack player were first announced in 1961 [Thorp, 1961, 1961a]. These strategies and all subsequent favorable player strategies are based on the fact that, as play proceeds, the player sees some or all of the used cards. Knowing that certain cards are missing from the pack, the player can in principle repeatedly recalculate his optimal strategy and his corresponding expectation.

In practice each card is assigned a point value as it is seen. By convention the point value is chosen to be positive if having the card out of the pack significantly favors the player and negative if it significantly favors the casino. The magnitude of the point value reflects the magnitude of the card's effect but is generally chosen to be a small integer for practical purposes. Then the cumulative point count is taken to be proportional to the player's expectation.

To a surprising degree, the player's best strategy and corresponding expectation depend only on the fractions of each type of card currently in the pack and only change slowly with the size of the pack. Thus the better systems "normalize" by dividing the cumulative point count by the total number of as yet unseen cards.

*An earlier version of this paper was presented to Annual Amer. Mathematical Society Meeting, Washington, D.C., January 1975.

The player's expectation with a full pack of n decks is approximately zero. It varies with the rules and with the number of packs. It falls off slowly as n increases. Typical values are +0.13% for $n = 1$ and -0.53% for $n = 4$ (computed by Julian Braun; the underscored digits may be off by one). Thus most point count systems are initialized at zero cumulative total for the full pack, and the normalized cumulative count is taken to indicate the change in player expectation from the value for the full pack. Most point count systems are further chosen to satisfy $\sum_{i=1}^{10} c_i n_i = 0$ where c_i is the point value of the ith rank (call Ace "rank 1", also observe 10, J, Q, K are indistinguishable under the rules of Blackjack so assign all of them rank 10).

The original point count systems, the prototypes for the many subsequent ones, were the five count [Thorp 1961], the ten count [Thorp 1961a], and the "ultimate strategy" [Thorp 1962]. An enormous amount of effort by many investigators has since been expended to improve upon these count systems (see attached partial bibliography). The current state of knowledge is shown in Table 1 (based on [Braun 1974] where further details are available).

The idea behind these point count systems is to assign point values to each card which are proportional to the observed effects of deleting a "small quantity" of that card. Table 2 (courtesy of Julian Braun, private correspondence) shows this for one deck and for four decks, under typical Las Vegas rules. One must compromise between simplicity (small integer values for the c_i) and accuracy. Thorp's ultimate strategy is a point count based on

Table 1. Braun's simulation of various point count systems. "Bet 1 to 4" means that 1 unit was bet except for the most advantageous x% of the situations, when 4 was bet. To compare systems, x was approximately the same in each case, 21(%).

	STRATEGY/SYSTEM	RESULTS OF SIMULATED DEALS-PLAYER'S ADVAN. Flat Bet	Bet 1 to 4	BASIC POINT COUNTS CARD 2	3	4	5	6	7	8	9	T	A	÷N?	SIDE ACE COUNT?	λ(C) 1 DECK	4 DECK
1	Basic Braun + -	.2%	1.4%	1	1	1	1	1	0	0	0	-1	-1	NO	NO	.963	.961
2	Braun + -	.7%	2.0%	1	1	1	1	1	0	0	0	-1	-1	YES	NO	.963	.961
3	Revere Pt. Ct.	.6%	2.1%	1	2	2	2	2	1	0	0	-2	-2	YES	NO	.979	.977
4	Revere Adv. + -	.5%	1.6% to 1.8%+	1	1	1	1	1	0	0	-1	-1	0	YES	YES	.885 (.968)	.884 (.968)
5	Revere Adv.Pt.Ct.-71	.6%	2.0%	2	3	3	4	3	2	0	-1	-3	-4	YES	NO	.994	.994
6	Revere Adv.Pt.Ct.-73	.8%	2.1% to 2.3%+	2	2	3	4	2	1	0	-2	-3	0	YES	YES	.924 (.995)	.922 (.997)
7	Thorp Ten Count	.7%	1.9%	4	4	4	4	4	4	4	4	-9	4	YES	NO	.720	.709
8	Hi-Opt	.8%	2.1% to 2.3%+	0	1	1	1	1	0	0	0	-1	0	YES	YES	.874 (.938)	.871 (.939)

Table 2 (J. Braun). Effect on player expectation E of removing one card of specified rank from a complete pack of one deck and four cards from four decks. Braun thinks actual one deck E are about .05% higher and changes in E are accurate to .01%. d is weighted sum of changes.

One deck $E_0 = 0.10\%$

Card i removed	A	2	3	4	5	6	7	8	9	10	d	d/13
E_i(%)	-.48	.50	.56	.70	.86	.56	.41	.12	-.06	-.39		
Change in E(%)	-.58	.40	.46	.60	.76	.46	.31	.02	-.16	-.49	.31	.024
u_i	-.604	.376	.436	.576	.736	.436	.286	-.004	-.184	-.514		

Four decks $E_0 = -0.532\%$

Cards i removed	A	2	3	4	5	6	7	8	9	10	d	d/13
E_i(%)	-1.130	-.147	-.081	.059	.236	-.078	-.239	-.525	-.714	-1.019		
change in E(%)	-.598	.385	.451	.591	.768	.454	.293	.007	-.182	-.487	.221	.017
u_i	-.615	.368	.434	.574	.751	.437	.276	-.010	-.199	-.504		

moderate integer values which fits quite closely the data then (1962) available. Until recently all the other count systems were simplifications of the "ultimate".

System 1 does not normalize by the number N of remaining cards. Thus the player need only compute and store the cumulative point count c. Normalization gives the improved results of system 2, but requires the added effort of computing and storing the additional number N and of computing c/N when decisions are to be made. In practice the player can estimate N by eye and use it with system 1 and get almost the results of system 2 with much less effort. Systems 2, 3, 5 and 7 are of c/N type.

Systems 4, 6 and 8 have the first new idea. They assign a point count of zero to the Ace for strategy purposes. This is consistent with the evidence: in most instances that have been examined, the optimal strategy seems to be relatively unaffected by changes in the fraction of Aces in the pack. However, the player's expectation is generally affected by Aces more than by any other card (Table 2). Therefore these systems keep a separate Ace count. Then the deviation of the fraction of Aces from the normal 1/13 is incorporated for calculating the player's expectation, hence for betting purposes.

Much of this paper parallels portions of the penetrating and important papers by Griffin [1975, 1975a]. For instance the $\lambda(C)$ values of Table 1, entries 1 and 2, 4, 6, 7, 8 correspond to the betting correlation numbers in [Griffin, 1975a] Table 7, rows 5, 4, 12, 8 and 1, respectively. The very close agreement is no accident. Both Griffin's correlation coefficient and our $\lambda(C)$ are the cosines of the angles between essentially the same pair of vectors.

2. Evaluating Point Count Systems: The "Probabilistic Gradient" Method. Suppose point count systems which are "closer" to the relative u_i values of Table 2 are likely to be "better". To test this we require a precise meaning for "better" and a precise measure of "closeness". We begin by basing the definition of "better" on the notions of probabilistic dominance, and of risk, used in mathematical finance [ ].

Definition 2.1. Let F and G be probability distribution functions. Then F probabilistically dominates G if $F(x) \leq G(x)$ for all x. If in addition $F(x_0) < G(x_0)$ for at least one x_0 then F strictly probabilistically dominates G. If F and G arise from random variables X and Y, respectively, or from probability measures μ and ν, respectively, then the defined terms apply to these pairs if they hold for F and G.

That F probabilistically dominates G is equivalent to $P(X \geq x) \geq P(Y \geq x)$ for all x. If X is the player expectation from point count system A and Y is the player expectation from system B, then this means that the chance of finding expectations of x or more is always as least as good using A as it is by using B. One can show that this means that a player following A has at least as great an expected return as B with "the same risk level".

However, probabilistic dominance is inadequate as a definition of "better" because the typical situation is that F is "spread out" more in both directions from the mean full deck expectation $E_0 \doteq 0$. Thus F dominates G for $x > E_0$ and G dominates F for $x < E_0$. In fact G is (to a good approximation) a convex contraction [Thorp, 1973] of F. More precisely, if E_F and E_G are the respective means of F and G, we will find $E_F \geqq E_G \geqq E_0$ with $Y - E_G$ a convex contraction (this is equivalent to the notion "less risky than" of portfolio theory; see [Hadar and Russell, Hanoch and Levy, Quirk and Saposnik, Thorp and Whitley] of $X - E_F$. Thus F is both "spread out more" than G and translated in the positive direction more. The reason why $E_F, E_G \geqq E_0$ is because E_0 is the expectation using the basic strategy and constant bets, equivalent to the full pack expectation. When (advantageous) counting systems are used, the strategy for playing hands is improved whenever the player has seen any cards other than the ones he and the dealer use on the first round. Since this generally happen with positive probability, we then have $E_F, E_G > E_0$.

Definition 2.2. Point count system A is <u>better than</u> system B if $E_F \geqq E_G$ and also $P(X \geqq x) \geqq P(Y \geqq x)$ for $x \geqq E_G$.

Typically count systems satisfy $E_F \geqq E_G \geqq E_0$ and $X - E_F \doteq a(Y - E_G)$, $a \geqq 1$ (a special case of convex contraction). These conditions imply A is better than B.

Assume that the betting systems $b(E)$ are numerican functions of the expectation E. Further assume $b(E)=1$ if $E\leqq 0$ and $b(E)\geqq 1$ if $E>0$. These are the ones generally considered. The popular fallacious systems such as the martingales (e.g. "doubling up"), and the La Bouchere [Thorp, 1962; Thorp, 1966; Wilson, 1965] which incorporate past results, are of no interest here.

Theorem 2.3. With the preceding notation and assumptions, if A is <u>better than</u> B, then for any betting system $b_B(E)$ based on the B point count, there is a betting system $b_A(E)$ based on the A point count such that the return R_A per unit bet by A (approximately) probabilistically dominates R_B. Further, R_A and R_B have approximately the same risk. In fact $R_A \doteq R_B + c$, where $c \geqq 0$.

Proof. If F and G are continuous, define b_A by $b_A\left(F^{-1}(G(E))\right) = b_B(E)$. Then note that the first unit of each bet has expectation E_A for A and E_B for B. The remainder of the bet is non-zero only if $E \geqq E_F$. Then for corresponding percentiles of their respective distributions, A places the same bets as B. But $F(E) \leqq G(E)$ if $E \geqq E_F$ so A has in each instance at least as great expectation, hence has at least as great expectation overall. Thus the total expected return to A is at least as large as for B. Also $R_A \geqq R_B$ per unit since the bets placed have the same distribution.

In reality F and G are not continuous; instead they are finite. But they may be arbitrarily closely approximated by continuous distributions so the result extends, with one qualification. If F or G is discontinuous extend the graphs of F and G by adding vertical segments at the discontinuity points so that the extensions $\tilde{F}$ and $\tilde{G}$ have inverses defined on $(0,1)$. Then for those E' such that G is discontinuous at E' or F is discontinuous at $F^{-1}(G(E'))$ it may be necessary to define $b_A(F^{-1}(G(E')))$ "probabilistically", so it is multiple-valued, each value occurring with specified probabilities.

To show that $R_A = R_B + c$, which implies the same risk, it suffices to assume that at each percentile level y for the distributions F and G we have the conditional distributions given y satisfying $F(x|y) = G(x-f(y)|y)$ where $f(y) \geq 0$. Since this only holds approximately in practice, we have $R_A \doteq R_B + c$.

Now we turn to the problem of measuring "closeness" of a given count to the "ultimate" strategy. We shall assume that point count strategies are of the form $C = (c_1, c_2, \cdots, c_{13})$ where c_1 is the value assigned for an Ace, $c_2, \cdots, c_9$ are the point counts for ranks 2 through 9, and $c_{10} = \cdots = c_{13}$ are the point counts for Tens, Jacks, Queens and Kings respectively. In practice these are lumped together and only ten point count values are specified. By writing C with 13 components we gain a symmetry which yields substantially simpler proofs. Note that C and aC, $a \neq 0$, are equivalent and will be identified.

Definition 2.4. If $\sum \Delta E_i = 0$ the <u>ultimate strategy</u> $U = (u_1, \cdots, u_{13})$ is the one given by $u_i = \Delta E_i$ where ΔE_i is the change in expectation from removing one ith card from the complete pack. If $d = \sum \Delta E_i \neq 0$ then U is given by $u_i - d/13$.

In Table 2, we have d for one deck is .024 and d for four decks is .017. Then the u_i rows are calculated in Table 2 from Definition 2.4.

It is tempting to think of U as representing to good approximation the direction of the gradient ∇E at $f_1 = \cdots = f_{13} = 1/13$ of the player's expectation $E(f_1, \cdots, f_{13})$ as a function of the fractions f_i of the cards from $i = 1$ to 13. Then we calculate $\lambda(C) \equiv C \cdot U/\|C\| \cdot \|U\|$, i.e. the projection of C in the ∇E direction. The numerator is the inner or scalar product and $\|C\| = \left(\sum c_i^{\,2}\right)^{\frac{1}{2}}$.

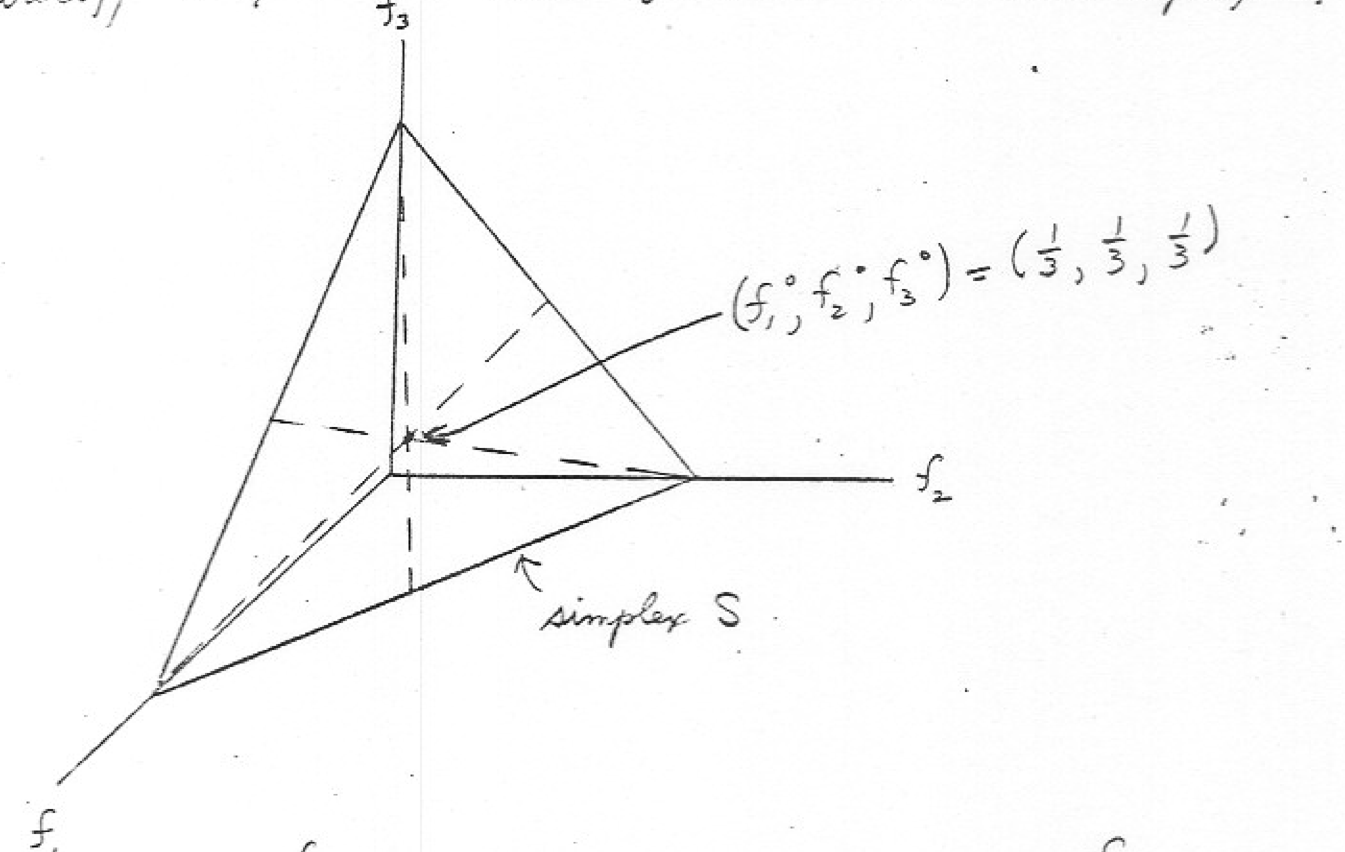
Next we claim that $\lambda(C)$ gives the approximate ratio of the spread of the C distribution F_C about E_C to the U distribution F_U about E_U. Then $\lambda(C)$ is the desired measure of closeness. In particular, for approximately the same risk per unit, and the same distribution of bet sizes, it would follow that $E(R_U) \doteq E(R_C)/\lambda(C)$. Then C_1 and C_2 are arbitrary strategies $E\left(R_{C_1}\right)/E\left(R_{C_2}\right) \doteq \lambda(C_1)/\lambda(C_2)$ for the same risk level and distribution of bet sizes. Thus the "power" of a strategy C is proportional to its $\lambda(C)$.

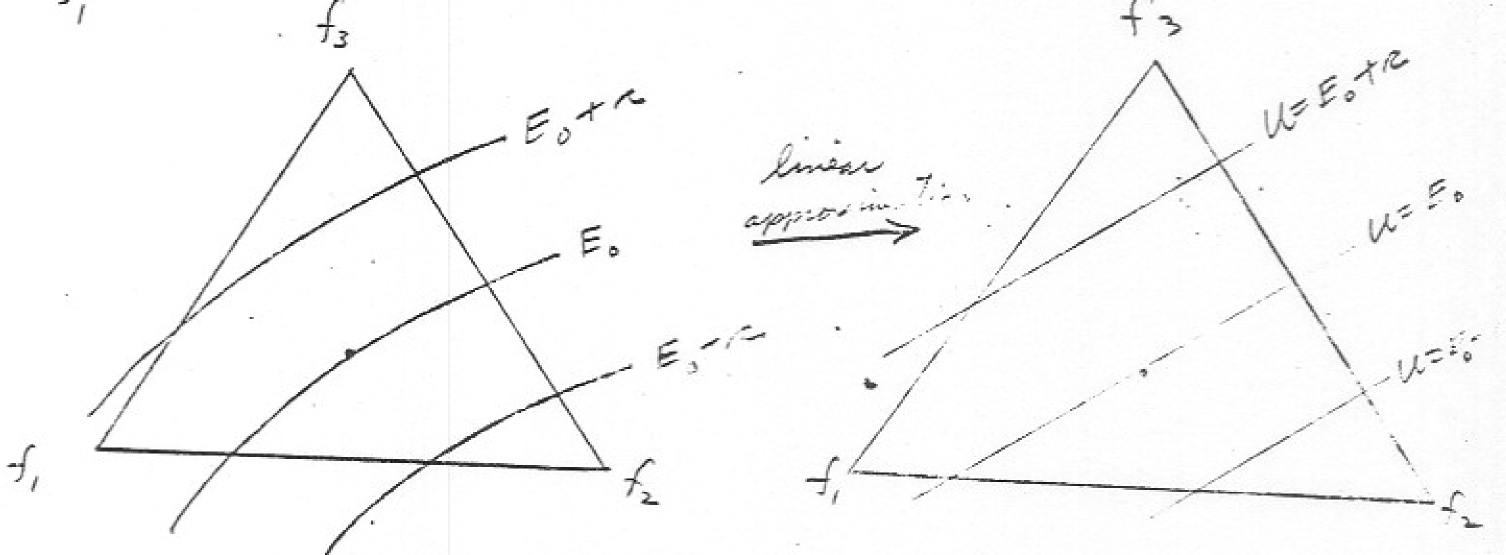
This conclusion is true but the argument must resolve two obstacles:

(1) In the preceding discussion we treated C, U, ∇E, etc. as though they were given in Cartesian coordinates when in fact they are not (see Figure 1).

(2) The probability distribution of $E(f_1, \cdots, f_{13})$ must be considered in reaching the conclusion and in general will invalidate it.

Figure 1. Illustration of $E(f_1, \ldots, f_n)$ and the U approximation when $n = 3$. The possible $(f_1, \ldots, f_n)$ satisfy $\sum f_i = 1$ and $f_i \geq 0$ so lie in the simplex S.





idealization from discrete to continuous: "curves of constant E"

Note further that both U and C are linear approximations to an in general curved "surface". Also in the real case the domain is a large finite subset of points of the possible $(f_1,\cdots,f_{13})$, each of positive probability. (The original discovery of winning blackjack systems [Thorp, 1961] was motivated by this model. First I introduced the $E(n_1,\cdots,n_{13})$ "surface", where n_i is the number of cards remaining of denomination i. Intuitive arguments "convinced" me that the E surface should have substantial deviations from E_0, the full deck expectation. The next step was to approximate by "the" $E(f_1,\cdots,f_{13})$ "surface", and then to "linearize" the problem by assuming that $E(f_1,\cdots,f_n) \doteq E_0 + \sum k_i \Delta f_i$, where $\Delta f_i = f_i - 1/13$.) Thus there is the approximation of a discrete problem by a continuous one. Nonetheless, we shall show:

Theorem 2.5. If the probability distribution of $(f_1,\cdots,f_{13})$ is approximately rotationally symmetric about $(1,\cdots,1)/13$ then the relative power of any point count system C is proportional to $\lambda(C) = C \cdot U/\|C\|\cdot\|U\|$. The powers of two count systems which exploit the count information equally (e.g. if one normalizes by the number of as yet unseen cards so does the other; if one carries a side Ace count for betting and sets the Ace equal to 0 for strategy, so does the other, etc.) are approximately proportional to their λ's.

Proof.

The Possibility of a New Kind of Winning Blackjack System.

The class of practical winning Blackjack strategies, known as card counting strategies, is well known. The player keeps track of the cards which have been used in play since the pack was last shuffled. He uses this information to vary his strategy for playing hands and to vary his bet, making "large" bets when his expectation is positive and "small" bets when it is negative.

A second class of winning Blackjack strategies has been suggested [Thorp, 1973] which uses the idea that in practice decks are not well shuffled, from a mathematical point of view. Here the player keeps track of the order in which the pack is stored as it is used. Then knowledge of the shuffling process, plus the order of appearance of cards when the pack is redealt, gives information about the distribution of the remaining cards. The player uses the non-randomness of the arrangement of unused cards to gain an advantage.

In this note we point out that there may be a third class of strategies. Here the player simply keeps track of the payoffs to him from previous hands (other than the first) from the current pack. He uses the information to determine his next bet.

Blackjack Bibliography

Abraham, Ralph, "Optimal Betting for Time Dependent Games." Preprint.

Braun, Julian H., "Comparing the Top Blackjack Systems," Gambling Quarterly, Fall/Winter 1974, pp. 22-23, 54-55.

Collver, Donald L., Scientific Blackjack and Complete Casino Guide. New York: Arco Publishing Co., Inc., 1966.

Einstein, Charles, How to Win at Blackjack. Galahad Books, 1968.

Epstein, Richard A., The Theory of Gambling and Statistical Logic. New York: Academic Press, Inc., 1967.

Gordon, Edward, Optimum Strategy in Blackjack-A New Analysis. Claremont Economic Papers, Claremont, CA, January 1973. Also in Gambling and Society, Charles Thomas Pub., Springfield, IL, 1975.

Griffin, Peter, "Rate of Gain in Player Expectation," Gambling and Society, Charles Thomas Publishing, Springfield, IL, 1975.

Griffin, Peter, "Use of Bivariate Normal Approximations to Evaluate Single Parameter Card Counting Systems in Blackjack." To appear, Proceedings of Second Annual Conference on Gambling, held June 1975a.

Hadar, J., and W.R. Russell, "Rules for Ordering Uncertain Prospects," American Economic Review, January 1969, pp. 25-34.

Hanoch, G., and H. Levy, "The Efficiency Analysis of Choices Involving Risk," Review of Economic Studies, 1969, pp. 335-346.

Heath, David, "Algorithms for Computation of Blackjack Strategies." To appear, Proceedings of Second Annual Conference on Gambling, held June 1975.

Noir, Jacques, Casino Holiday. Berkeley: Oxford Street Press, 1968.

Palais, Richard, "Computer Simulation Studies of Blackjack Systems." Available from the author.

Quirk, J.P., and R. Saposnik, "Admissibility and Measurable Utility Functions," Review of Economic Studies, 1962, pp. 140-146.

Revere, Lawrence, Playing Blackjack as a Business. Las Vegas: Paul Mann Publishing Co., 1971; Lyle Stuart, 1973.

Thorp, Edward O., "Fortune's Formula: A Winning Strategy for the Game of Blackjack," Notices, A.M.S., January 1961, page

Thorp, Edward O., "A Favorable Strategy for Twenty One," Proceedings of the National Academy of Sciences, 47, 1961, pp. 110-112.

Thorp, Edward O., Beat the Dealer. New York: Random House, 1962. Revised, 1966.

Thorp, Edward O., Elementary Probability. New York: John Wiley and Sons, 1966.

Thorp, Edward O., "Optimal Gambling Systems for Favorable Games," Review of the International Statistical Institute, 37:3, 1969, pp. 273-293.

Thorp, E.O., and W.E. Walden, "The Fundamental Theorem of Card Counting," International Journal of Game Theory, Vol. 2, 1973, Issue 2.

Thorp, E.O., and R.J. Whitley, "Concave Utilities are Distinguished by Their Optimal Strategies," Colloquia Mathematica Societatis Janos Bolyai, 9. European Meeting of Statisticians, Budapest (Hungary), 1972.

Wilson, Allan N., The Casino Gambler's Guide. New York: Harper and Row, 1965.