In *The Journal of Philosophy* (1958), **55** (5) February: 212–216.

Theory of Games as a Tool for the Moral Philosopher. An Inaugural Lecture Delivered in Cambridge on 2 December 1954. R. B. BRAITHWAITE. Cambridge: Cambridge University Press; [New York: Cambridge University Press] 1955. 76 pp. $1.25.

This little book, which is the printed version of Professor Braithwaite's Inaugural Lecture, is in the great tradition of Bernoulli, Bentham, Sidgwick, and Edgeworth. No doubt any contemporary philosopher who mentions a calculus of utility or any other sort of calculus in connection with moral problems immediately thinks it necessary to assume a somewhat defensive attitude. But Braithwaite justly observes on the first page that he is sensitive to another sort of criticism, namely that academic moral philosophers never deal with the solution of actual moral problems, but are entirely concerned with the epistemological status of moral concepts and the logic of moral language.

Braithwaite's moral problem is the important one of offering two people advice on how they should behave in a situation where they have both mutual and conflicting interests. In the language of game theory he is proposing a principle for solving two-person coöperative games. When his endeavor is so described, many philosophers would be inclined to retitle his work "A moral philosopher dabbles in game theory,'' but this would be a gross mistake. The order of development is, rather, that a situation involving two persons in a moral dilemma is presented. The moral problem is to resolve this dilemma in a just manner. And concepts from game theory are brought in as tools to aid in its resolution.

As a literary device Braithwaite expounds his analysis for the case of two bachelor musicians, Luke and Matthew, who live in adjoining flats and are faced with the problem of equitably deciding when each is to play his instrument. The first stumbling block is the measurement of their individual preferences for the four possibilities of: both playing, neither playing, Luke playing alone, Matthew playing alone. Hurdling this obstacle is not a point of emphasis with Braithwaite. He adopts the probability mixtures approach to the measurement of preference or utility originally due to F. P. Ramsey (1926) and now in much prominence since its use by von Neumann and Morgenstern in their treatise (1943) on game theory. To those philosophers skeptical about the feasibility even in principle of such measurement, it may be retorted that a growing experimental literature shows that in restricted situations individuals actually behave so as to satisfy axioms which guarantee appropriate numerical assignments to their preferences.

Even when individual numerical preference scales are granted, rational methods are hardly apparent for resolving conflicts between individuals. When the conflict involves only two individuals and their aims are diametrically opposed, that is, the situation is wholly competitive (a zero-sum, two-person game), the minimax principle of von Neumann is highly satisfactory. But few conflicts of interest are so total in character, and since 1943 a literature has developed around the problem of equitably resolving such conflicts both when opportunities for bargaining and arbitration exist (coöperative games) and when no such opportunities exist (non-coöperative games). Braithwaite's situation is of the coöperative type, and his proposed solution builds on previous work of Nash and Raiffa, particularly the latter. In his article ''Arbitration Schemes for Generalized Two-Person Games'' (*Annals of Mathematics Studies*, No. 28, Princeton, 1953), Raiffa offers an arbitration scheme which resolves fairly satisfactorily the dilemma of Braithwaite's bachelor musicians, *provided* a meaningful interpersonal comparison of their individual utility functions may be made. Roughly speaking, the arbitration scheme consists of defining lines of constant relative advantage in the space of possible outcomes of arbitration. For each strategy pair—one strategy for Luke and one for Matthew—a certain relative advantage obtains to a given player, say Luke. Using relative advantage to Luke as a payoff function, we obtain a non-coöperative two-person, zero-sum game. The von Neumann minimax solution v of this game is found. This solution v lies on a line V of constant relative advantage. The arbitration scheme is to recommend that the two bachelors now coöperate to increase their payoffs as much as possible by moving to the most favorable point of V. (Precise definition of "most favorable point" is not difficult; in economic terms it is a Pareto optimal point in the space of possible outcomes.)

Braithwaite obtains Raiffa's results by some elegant arguments which exploit the geometry of the space of outcomes. More importantly, Braithwaite extends Raiffa's analysis by proposing a natural common unit of utility and eliminates, if acceptable, a major difficulty for Raiffa, namely, the interpersonal comparison of utilities. The proposal, Braithwaite says, "rests upon the discovery that the logic of the general collaboration situation is isomorphic with the geometry of a parabola regarded as an envelope of lines, . . . the parabola being uniquely determined by the four pure strategies available to the collaborators" (p. 27). Without going into geometrical details, the proposal is to assume that Luke and Matthew each benefit equally by the change from a prudential strategy (maximin strategy) to a counter-prudential strategy (minimax) while the other holds to a prudential strategy. Unfortunately Braithwaite's defense of this assumption rests almost entirely upon relatively intricate geometrical arguments, primarily concerning questions of symmetry. I say "unfortunately" for it is not clear that his bachelors or other parties to an arbitration parley would be much influenced by geometrical considerations in deciding to accept or reject the arbitrator's advice.

It would be worthwhile to attempt to find principles which are more behavioristic to bolster Braithwaite's assumption. Any suggestion concerning "natural" units in the area of subjective measurement must be viewed with great skepticism unless it can be shown to be a logical consequence of relatively simple, elementary principles which do not baldly postulate the unit. (Such principles can be found, for instance, to justify the natural zero for measurements of subjective probability, but not so easily a natural unit.)

These last remarks are not meant to belittle Braithwaite's results concerning a notoriously difficult problem. Interpersonal comparison of preferences or utilities has been the *bête noire* of welfare economics for at least two decades. It is most heartening to find a philosopher who is willing and able to tackle this problem, and to see its relevance to moral philosophy.

Since the moral problem posed by Braithwaite and the methods used to expound a solution are scarcely dominant in current ethical writings, I would like in conclusion to make two observations. The first is that when a moral dilemma may be formulated with any degree of precision the stock tools of philosophical analysis are of little use. As Braithwaite puts it, what he does in his book is give an explication or rational reconstruction of the concept of fair play as applied to a two-person arbitration situation. The concepts he uses cannot be derived from the logical study of the language of morals. It would be as foolish to attempt to derive an arbitration solution from the language ordinarily available to Luke and Matthew as it would be to attempt to derive the theory of probability from the professional talk of gamblers.

To be sure, the standard objection to Braithwaite's problem is that it is unrealistic to think any actual moral dilemma can be stated with such precision. When I told one philosopher of Oxonian persuasion about the bachelor musicians, he laughed and said, "Whoever heard of such absurd computations." No doubt an example dealing with arbitration between a labor union and a corporation would have made the charge of absurdity less patent, but the particular example used to develop the ideas is no more essential than the character of the particular examples used in a textbook of mathematics. In learning mechanics or the differential and integral calculus, it is understood that one does not use these disciplines to decide how large a glass of water to drink on a hot afternoon, or how fast to drive a car on an icy road, although such topics might be suitable for textbook problems. Admittedly welfare economics and the theory of games are not yet powerful applied disciplines like mechanics, but substantial progress has been made, and the appraisal of that progress should not be initially biased by taking too literally any particular example illustrating the theory.

My second observation is that Braithwaite is surely on the right track in his direct and unashamed use of mathematics in analyzing a moral problem. For those who hold that ethics is the logical study of the language of morals, it is natural to pursue the various deontic logics of imperatives, obligations, etc. It is my own conviction that the development of such logics has lead almost entirely to results which are mathematically trivial and philosophically uninteresting. Anyone who has made a serious attempt to apply a formal system of logic to an empirical or philosophical problem soon realizes that logic is a technically weak instrument compared to classical mathematical analysis and geometry. It would be a pity if formal developments in ethics during the next few decades were restricted to formal logics when so many relevant mathematical concepts are now at hand from game theory, welfare economics, and statistical decision theory. For, in my own opinion and Braithwaite's too I gather, the fundamental problems of moral philosophy are conceptual problems of evaluation and decision, individual and social, not problems of language.

PATRICK SUPPES

STANFORD UNIVERSITY