Faculty Submissions   **Philosophy**

9-2003

The Moral Poker Face: Games, Deception, and the Morality of Bluffing

James McBain
Pittsburg State University, jmcbain@pittstate.edu

Recommended Citation

THE MORAL POKER FACE:
GAMES, DECEPTION, AND THE MORALITY OF BLUFFING

James McBain
Kansas State University, KS

Bluffing, when you get down to it, is nothing more than a type of deception. But, despite its morally questionable foundation, it is not only permissible in certain contexts, but sometimes encouraged and/or required. For example, when I am playing poker and have a fairly poor hand (say, a mere pair of 10's), and bluff my way to winning the hand, I not only win without question, but am praised for my deception. In this context – that of playing poker – bluffing is morally permissible. But, the question remains as to whether it is permissible to bluff in other contexts – particularly normal, everyday situations.[1] In this paper, I will look at the argument given by László Mérő – an argument based in game theory and Kantian ethics – to the end that bluffing is morally permissible in everyday contexts. I will argue that Mérő's argument is mistaken on two grounds. First, it includes an epistemic feature (i.e., knowledge that bluffing is part of the game) that is lacking in everyday contexts. Second, even if we add a proviso to solve this epistemic problem, the resulting strategies fail to guarantee an equilibrium state. Thus, I hope to show that Mérő fails in his attempt to justify the use of bluffing in everyday contexts.

BLUFFING – A WORKING DEFINITION

Before addressing the morality of bluffing, it is necessary to provide a working definition of bluffing.[2] Bluffing, in essence, is a form of deception. When one bluffs, whether in poker or in other contexts, she attempts to impress, deter, or mislead her opponents by a false display of confidence. For example, in poker, one can bluff in two ways: first, one may give a false impression of strength when she is actually in a weak situation, or second, one may give a false impression of weakness when she is actually in a strong situation (Von Neumann and Morgenstern, 189). But bluffing is not confined to these types of betting situations. Bluffing can also involve "trash-talking," show-boating, and deceptive facial expressions.

What is important is that the bluffer does not make any untrue statement about the object of deception. To make an untrue statement would be to lie. Consider the following situation: Jones is interviewing for a job. The company offers Jones the position. In an attempt to "sweeten the deal," Jones tells the interviewing committee, "Well...I have some other things to consider before I sign with your company, may I inform you of my decision on Monday?" Further suppose that there is nothing else for Jones to consider. She is going to take the position whether the company offers more money or a better package. In this case, Jones has not said anything overtly untrue (we could say that there are trivial things for Jones to consider, say, how far it is to drive or which is the best driving route to the new job). Jones is playing a strategy where she is intending to deceive the interviewers into believing that she has another job offer and that they should up the ante as to get her to sign with them. There is no direct information about the state of Jones' situation being conveyed; hence Jones is not lying, merely bluffing. Consider a second case, Jones is in the same situation, but when offered the position, Jones says, "Well...I have a couple of other job offers I have to consider before I can give you a definite answer, may I let you know on Monday?" In this second case, Jones' intention is the same, but here she is stating something that is untrue. Here she is lying. So, in order to bluff, one must not state anything untrue. The behavior has no direct informational content and as such it is not an actual lie.

Finally, the motivations in bluffing and lying are different. When one bluffs, she is attempting to deceive, but does not mind if her bluff is called.[3] When someone's bluff is called in one round, the next time she may gamble for bigger stakes since the opponents will expect a bluff. In this way, bluffing is a

long-term strategy. Lying is just the opposite. The liar does not want her lie called. And if her lie is called, then no one will want to play again with her. In this way, lying is an attempt for short-term gain with the possibility of long-term loss. So, a bluff, as I am taking it, is a form of deceptive (verbal or non-verbal) behavior where one intends to reach an advantageous position by making her opponents believe that there are other factors involved in the game that are not actually involved.

THE PERMISSIBILITY OF BLUFFING IN GAMES

One way to argue for the permissibility of bluffing that has been much discussed is to draw an analogy between the context in question and a game context (e.g., poker). Albert Z. Carr argues for the permissibility of bluffing in business situations on the grounds that business contexts are simply games and, like poker, in games we are allowed to bluff. It is part of poker that there is bluffing allowed, and as such, in business bluffing should be allowed.[4] The problem with generalizing this argument to everyday contexts, is that it is not clear as to whether bluffing is permissible since the scope for bluffing is neither agreed to nor well-understood by everyone involved.[5] It is simply a matter of not knowing in which contexts the analogy would hold. So, it would seem that we cannot merely say that since bluffing in poker is permissible, bluffing in other game-contexts (which would include everyday situations) is permissible. Without knowledge of the rules and permitted behaviors of the game context, the argument by analogy fails to justify bluffing as permissible in everyday contexts.

Another, more interesting, argument concerns the bluffing strategy itself. László Mérő argues that it is permissible to bluff in everyday contexts on the grounds that no equilibrium state can be maintained in the long run without such behavior. His argument is an unlikely marriage of game theory and Kantian ethics. Mérő attempts to show that bluffing is necessarily part of optimal (mixed) strategies in an equilibrium state; and, we can find maxims based on those strategies that pass the categorical imperative. Thus, he attempts to show that bluffing is not an ethically condemnable act.

Let me attempt to cash out Mérő's argument by addressing each part separately. First, Mérő argues that bluffing is part of any type of competition and that no equilibrium can be reached without such behavior. That is, bluffing behavior can lead to an equilibrium state that is optimal for all players and this state is one where no one's interests would dictate changing their strategy (Mérő, 68). Moreover, certain types of (mixed) strategies necessarily contain bluffs.[6] To illustrate this, let us look at the poker model that Mérő presents.

Suppose there is a two player game (between X and Y) where X is always the challenger and Y is the challenged. The game is this: X rolls a normal, six-sided die, and if it shows a six, then he wins, if it is not a six, he loses (Mérő, 70-72).[7] Consider various strategies X might adopt. Suppose X never bluffs. Here, X will run a deficit in the long run. In the long run, Y will always believe X and so when X wins, she wins $30 and when she loses, loses $10. But she will win only 1 out of every 6 rolls hence losing $50 for every $30 she wins. So, this is not an optimal strategy. Suppose X bluffs a lot. This strategy will result in a lack of profit. If X always raises, Y will no longer believe X and will not fold. Thus, for every six rolls Y will win a total of $300 and X will win $80 once. Suppose that X adopts the following mixed strategy – whenever X throws a six, she raises; otherwise, she bluffs with a probability of 1/9.[8] This strategy results in a modest, but certain long-term profit (approximately, $20 a roll).[9] If X plays this strategy without deviation, Y can do nothing but lose money. It does not matter whether Y folds, X will continue to average $20 a round. Due to this, these strategies are in equilibrium (i.e., neither player has an incentive to switch strategies since neither can do better by switching strategies).[10] Thus, for this simple game, the only strategies that can reach equilibrium contain bluffing as part of the strategy.

It is acknowledged that the "bluff" is not part of the mathematics of the game sketched above. The bluff is in terms of different possible moves. That is, bluffing occurs any time a move is made with higher probability than the optimal (mixed) strategy would dictate. Mérő's point is that in mixed strategies ending in equilibrium states, the equilibrium state can be reached only if the players are

occasionally bluffing.[11] Thus, as Mérő states, "bluffs must appear in every kind of competition, for no equilibrium can be maintained in the long run without them" (Mérő, 78).

The Kantian twist Mérő gives his position is that there are going to be situations in which we can act upon a maxim that can satisfy the categorical imperative only if it contains bluffing behavior. More specifically, any time the players of the game play an optimal mixed strategy, they will not violate the categorical imperative. Since an optimal state is going to be one in which no one will want to change strategies and no one can do any better off, if everyone acted according to their optimal mixed strategy, there is no inconsistency. Thus, as long as people play according to optimal mixed strategies, the action dictated by the maxim is morally permissible.

Combining the two points, we see Mérő's conclusion that bluffing is morally permissible. Since any competitive situation in equilibrium contains an optimal (mixed) strategy that includes bluffing behavior, those strategies do not violate the categorical imperative. This is due to there being no inconsistency in a state where everyone played according to the optimal (mixed) strategy. Thus, any optimal (mixed) strategy that contains bluffing behavior does not violate the categorical imperative and, hence, is morally permissible.[12]

AN INITIAL DIFFICULTY WITH MÉRŐ'S ARGUMENT

There is a problem with Mérő's argument is similar to the problem with the argument by analogy previously mentioned. The problem there was that the scope of bluffing is not sufficiently well-defined in order for the analogy to hold in everyday contexts. That is, in games like poker, bluffing is known by all parties to be part of the game itself. In everyday contexts, there is no such knowledge. Mérő faces a similar problem. The poker model he uses to provide the basis of his argument (to show that bluffing is a necessary part of any optimal mixed strategy) relies on the fact that both players know that bluffing is part of the game. This knowledge allows the mixed strategy containing bluffing to pass the categorical imperative.

Mérő's model fails once we rework the model to allow for the fact that players may or may not know that bluffing is part of an acceptable strategy. Let us suppose that the strategy being played is still optimal. Would it still pass the categorical imperative? While it may still be possible to conceive of a world in which everyone played an optimal mixed strategy, it is questionable that it is a maxim that respects the dignity of the players involved. That is, the rules of the game are such that we cannot determine whether all the players would rationally consent. Since the fact that bluffing is permitted is not known, we cannot say that all players are fully informed and hence they are not being respected as persons when playing (remembering that the "game" in question is an everyday context, not a rule-stipulated game like poker). In this way, the strategy, while still being optimal, fails to pass the categorical imperative. Thus, playing the strategy is immoral according to one formulation of the categorical imperative.

AN EASY FIX?

One might object that I am being too literal in my assessment of Mérő's argument. It might be said that the lack of knowledge that the agents have could easily be inserted into the strategy and thereby avoid the difficulty discussed. For instance, we could supplement Mérő's model with the conditions on bluffing expressed by Thomas Carson. According to Carson, bluffing is permissible when one has good reason to believe that the other player(s) are bluffing.[13] In other words, Carson sees the permissibility of bluffing as a sort of self-defense. That is, since I will be harmed as a result of the other player's bluffing, it is permissible for me to bluff as well since it is in my self-interest and morally permissible for me to defend myself against possible harm (Carson, 326-327).

This would seem to solve the initial worry about the lack of knowledge that the players have in everyday contexts. For instance, if I am looking to buy a car and I go to the dealer and she informs me

that the "asking price" is $20,000, and I know that she is a person who usually bluffs in the beginning of any negotiation, then it is permissible for me to bluff as well in the negotiation. So, Mérő's model has been changed, but not significantly. Any optimal mixed strategy that contains bluffing as a part is permissible given that one has good reason to believe that the players have bluffed in the game and/or will bluff in the game.[14] In this way, the new strategy has no problem passing the categorical imperative. Since all players would rationally consent to playing the game (they would have sufficient knowledge of the permissible strategies in the situation) and everyone would agree with the strategies that are in equilibrium (they are optimal), the fact that bluffing is part of the strategies does not violate the categorical imperative.

A PROBLEM WITH THE STRATEGIES

Though this epistemic proviso solves the problem of the strategies passing the categorical imperative, it is not clear that we have a state of equilibrium anymore. In other words, is it the case that the players have no incentive to switch strategies? In order to answer this question, it must be remembered that strategies are in equilibrium when neither player prefers to change unilaterally to another strategy (Resnik, 129).[15] That is, neither player can do any better by switching strategies. Now, when a player enters into a game (i.e., everyday context) she must assess whether she knows or has good reason to believe that her opponent has bluffed concerning her initial bargaining position or will bluff further into the game. Suppose that the player in question does not know that bluffing is acceptable, but has some reason to believe that her opponent will bluff (though it may not be that she has good reason to believe). In this case, she must decide whether she will add bluffing to her strategy. If she does and it is the case that bluffing is permissible, then there is no problem. Her strategy may be optimal and the strategies being played may reach equilibrium. If she does and it is not the case that bluffing is part of the game, the consequences may be that she is seen as intentionally being deceptive and will be excluded from future games. The problem here is that there is no way in which to determine whether bluffing is permissible by all parties involved.[16] And this being the case, it is not clear that the players have no incentive to switch strategies. Since it is not the case that there is no incentive to switch strategies, we cannot say that we will definitely be in equilibrium with bluffing as part of the strategy.

If this is correct, then Mérő faces a serious problem. If we take his model as is, then it would seem that there is an epistemic problem that fails to show that strategies that contain bluffing pass the categorical imperative. Yet, if we attempt to solve this problem with an epistemic proviso stating that it is permissible once the players have good reason to believe that the other players have bluffed or will bluff, then we solve the problem. But this solution is untenable. If we add the epistemic proviso, then it is not clear that the players will have no incentive to switch strategies. It may be the case that a player has an incentive to switch to a strategy that does not include bluffing as part of the mixed strategy since the long-term consequences of bluffing in an everyday context are too great. Thus, Mérő's argument for the permissibility of bluffing in everyday contexts fails.

CONCLUDING REMARKS

What I hope to have shown here is that Mérő's argument fails to support his conclusion that bluffing is an essential aspect of any optimal mixed strategy in everyday contexts. This is not to say that it is impossible to justify bluffing as a permissible strategy in everyday contexts. Rather, I maintain that if we are to justify the use of bluffing in everyday contexts, it is not the case that we can do so by appealing to the game theoretical/Kantian twist that Mérő endorses.[17]

REFERENCES

Carr, Albert Z. 1968. *Business as a Game*. New York: The New American Library, Inc.

Carson, Thomas. 1993. "Second Thoughts about Bluffing", *Business Ethics Quarterly*, 3(4): 317-341.

Koehn, Daryl. 1997. "Business and Game-Playing: The False Analogy", *The Journal of Business Ethics*, 16: 1447-1452.

Mérő, László. 1998. *Moral Calculations: Game Theory, Logic, and Human Frailty*. New York: Springer – Verlag New York, Inc.

Resnik, Michael D. 1987. *Choices: An Introduction to Decision Theory*. Minneapolis: University of Minnesota Press.

Von Nuemann, John and Oskar Morgenstern. 1953. *Theory of Games and Economic Behavior, Third Edition*. Princeton: Princeton University Press.

ENDNOTES

[1] While most of the literature on bluffing is in the realm of business ethics; here, I am not concerned specifically with business ethics. I will address business-related situations when necessary, but I intend my conclusion to address the morality of bluffing in everyday contexts.

[2] The account of bluffing developed here is based in part on Mérő (1998), Von Neumann and Morgenstern (1953), and Carr (1968). I do not take this to be the final word on the matter, merely a working account in order to assess arguments given in defense of the permissibility of bluffing.

[3] This is to say that she accepts that her bluff may be called. There are cases, such as war, where the stakes of having one's bluff called are too great and hence one minds greatly if her bluff is called. These cases aside, it is acknowledged that having one's bluff called is a legitimate possibility in engaging in the strategy.

[4] This is basically the argument given in *Business as a Game* (p. 173-201). It should be noted that Carr does not attempt to use this argument to justify bluffing in everyday contexts, just business contexts. I am merely addressing this argument as one possible way in which to justify bluffing in everyday contexts.

[5] This is a variation on the point made by Daryl Koehn (1997). She argues that the analogy between business and poker breaks down for various reasons, one of which being that the scope of bluffing in poker is narrow and well-defined whereas in business it is not. Though our objections are similar, she bases her objection on the fact that the players in business do not know the far reaching consequences of their bluff whereas in poker they do. My argument here is that in everyday contexts, the players do not know that bluffing is part of the game at all (and if they do, the rules for bluffing are not well-defined, nor agreed to by everyone playing).

[6] A mixed strategy is one where the player assigns a probability to each possible move in the game and then decides on the basis of these probabilities which to choose. The choice itself is dictated by chance, but the probabilities for each decision are not necessarily equal. This is to be contrasted with a pure strategy where one's decisions are dictated by some principle (that is, they are nonrandomized).

[7] The other rules of the game are as follows: (1) At the beginning of each round X bets $10 and Y bets $30. (2) X casts the die so that Y cannot see the result. (3) Having looked at the result, X can either fold or raise the bet. If X folds, Y takes all the money. If X raises, he has to put an additional $50 on the table. (4) If X raises, Y can either fold or not. If Y folds, then X takes all the money. If Y does not fold, then Y must put down $50. (5) If X raised and Y did not fold, then X must show the die. If the die shows a six, X wins all the money. If X was bluffing, Y wins all the money (70-71)