polish 4(184)'13
sociological
review
ISSN 1231 – 1413

MAREK M. KAMINSKI
University of California

GRZEGORZ LISSOWSKI
University of Warsaw

PIOTR SWISTAK
University of Maryland

Formal Theory and Value Judgments[1]

Abstract: When we assume that a survey reveals respondents' true attitudes we tacitly assume that the subjects understood what we are asking them about and that they had no incentive to be untruthful. In typical studies none of the two assumptions holds. Subjects are asked questions that use undefined terms and they are asked about issues they have no incentive to answer truthfully. Here we argue that a way to solve the two problems lies in constructing a formal theory of an attitude in such a way that an attitude can be derived from the answers yet when answering the questions subjects cannot possibly know that their responses reveal anything about their attitude and, hence, they have no incentive to answer insincerely. We briefly discuss a study that has the desirable properties necessary for the proper design.

Keywords: formal theory, experiments, value judgments, rationality.

Perhaps the most famous, now classical, finding about the American breed of homo politicus is the absence of politicus in the homo. Time and time again, voters in the US showed no evidence of clear ideological thinking (only about 2.5% of the electorate showed such evidence in the famed Campbell et al. 1960 study), and when they did, patterns over time suggested that their attitudes might have been random (Converse 1964). A gruesome picture indeed.[2] (To be fair, we should add that voters in other countries do not really look much better (Klingemann 1979).)

Yet the image of John Doe, the voter, gets considerably grimmer when we look at his strategic abilities. Data from many experiments on gaming paint a picture of an unsophisticated player wandering somewhere off the equilibrium path and rarely, if ever, converging on a maximizing choice (e.g., Rapoport, Guyer and Gordon 1976, Davis and Holt 1993). Experiments testing rationality in decision making under uncertainty (Kahneman, Slovic, Tversky 1982) only add to the murky image. The Allais Paradox, gambler's fallacies, effects of framing, and many other paradoxical departures from the principles of rationality stand as a rule rather than an exception (Piattelli-Palmarini 1994). In sum John Doe comes through as both ignorant and

[1] The support of the Center for the Study of Democracy (Kaminski) is acknowledged.

[2] Despite some positive change that has taken place over the years this early vision of an American voter has left a lasting imprint on our perception of homo politicus (cf., Niemi and Weisberg 1984).

irrational—"the truly unsophisticated cretin" to use the more exact phrasing of Niemi and Weisberg (1984: 322). The question is: Is he indeed?

The answer is of obvious importance for much of the empirical research in political science and sociology. The assumption of rationality is not only theoretically central to any social science but also critical for many empirical studies like those based on the spatial theory of voting, for instance (cf., Enelow and Hinich 1994). The concept of ideology carries similar weight in both theoretical and empirical considerations (cf., Hinich and Munger 1994). Ideology is central, for instance, for the social-psychological tradition of the Michigan school, whether it is in the form of a directly measurable property, like in the National Election Studies, or as a latent property that can be statistically extracted from other variables, like in Inglehart (1990). Yet, if we deal with a non-ideological and non-rational voter, then much of what we think we discover, be it through factor-analyzing issue positions or by asking people for ideological self-identifications, may be meaningless. Homo politicus and homo sociologicus who randomly change from liberal to conservative would destroy a good part of the mainstream political science and sociology.

Thus, whenever research results point to random, inconsistent, or irrational choices, potential implications for any social science can be profound. We have to be reasonably certain that these findings are not artifacts of a bad research design.

In this paper we propose to reflect on how to construct valid indicators in measuring properties such as rationality or ideology. In particular, we focus on how to measure attitudes related to value judgments (like ideological attitudes), when subjects may have little or no incentive to reveal their true positions (say, when surveyed by a state agency in an oppressive totalitarian regime).[3] We formulate two general principles that a construction of valid indicators should follow and argue that any such construction requires that we have a formal theory related to a phenomenon studied.

The first principle follows from an obvious observation that people can only have a shared understanding of what they are asked about if the concepts used in the question are properly defined. But well defined concepts can only be derived from a properly constructed theory—a formal theory. It is reasonable, for instance, to expect that when asked about their age everyone will have the same understanding of what they are asked about since aged is expressed by integers and integers are properly defined in arithmetic. It is not reasonable, however, to expect that when asked how liberal they are on a seven point scale subjects would have a shared understanding of what they are asked about even if we graciously assume that they have an understanding of what liberal means. This, naturally, is a consequence of "liberal" being an ill-defined concept.

The second principle follows from an observation that when we ask people about a normative matter we cannot assume they will give us a sincere answer. This means that true attitudes can only be discovered if we can derive them from questions that are not perceived as related to these attitudes. This, we argue, is only possible if there is a fairly complex deductive relation between questions and an attitude, a relation

[3] Incidentally, the consequences of bad survey design in a totalitarian regime can be catastrophic for the regime itself (Kaminski 1999).

that cannot be possibly recognized by subjects. This, again, requires that the attitude is defined in a formal theory where it can be deductively linked to a set of simple properties.

Our general methodological discussion is aided with two examples of research design that overcome, in our opinion, these two common problems with validity. We briefly sketch both studies and present some of their results. These results, finally, address the issue brought up in the opening paragraph. The picture that emerges from the two studies differs from the received view to a degree that is quite astonishing. People, as we will see, turn out to be remarkably consistent and rational in their choices.[4]

The First Lesson—Rational Choice in Two-Person Zero-Sum Games

As routine is the worst enemy of progress, revisiting old and seemingly well settled issues should never be treated lightly. At least such seems to be the lesson from the study of Barry O'Neill (O'Neill 1987). O'Neill, apparently suspicious of the past failures of the minimax theory for two-person zero-sum games in predicting choices in gaming experiments, assumed that the problem must lie with the design of experiments rather than with rationality of the subjects. To test this, O'Neill used the game of Table 1. (The game has an elegant justification in terms of theoretical simplicity of its design.)

Table 1

Payoffs to the Row Player in the O'Neill's Experiment

	Joker	Ace	Two	Three
Joker	5	−5	−5	−5
Ace	−5	−5	5	5
Two	−5	5	−5	5
Three	−5	5	5	−5

Even though the game seems sufficiently simple to be understood in the matrix form of Table 1, O'Neill assumed safely that subjects would be more likely to comprehend the payoff structure and the strategic aspects of the game if it were presented as the following card game. The game is played by two players. Each of the two players gets four cards—three number cards (ace, two and three) and a joker. They each play one card simultaneously. The Row player wins 5 cents if there is a match of jokers or a mismatch of number cards; otherwise he loses 5 cents.

The minimax solution prescribes that each player uses a mixed strategy (0.4,0.2,0.2,0.2). The solution is invariant with respect to players' utility functions

[4] For recent surveys of the rapidly growing literature on rational choice approach see Kronenberg and Kalter (2012) and Wittek, Snijders and Nee (2013).

as long as they prefer to win than to lose. Playing minimax makes the Row player win 40% of the time and gives him an expected payoff of −1 per play. So much for the predictions of game theory. Now the results of an experiment. In O'Neill's experiment with 25 pairs of subjects playing 105 iterations of the game, the first strategy of the Row player was played on average 0.396 times compared to the predicted 0.400 (a difference of 1%!) and the proportions of wins by each player were (0.401,0.599) compared to the (0.400,0.600) predicted by the minimax.[5] These are striking results—by any measure.

So what is John Doe—an irrational player of the early experiments, or the remarkably rational player of the O'Neill's study?

Aumann's opinion in that matter is that "(…) experimentation in rational social science is subject to peculiar pitfalls, of which early experimenters appeared unaware, and which indeed mar many modern experiments as well. These have to do with the motivation of the subjects and their understanding of the situation" (Aumann 1987: 7).

Since Aumann has written this we know considerably more and all that we know seems to confirm Aumann's opinion. O'Neill's study has revived interest in experimenting with simple zero-sum games and minimax solutions. Rather quickly the early experiments pointing to complete lack of rationality in players have been offset by experiments pointing in precisely the opposite direction. Rapoport and Boebel (1992) have, essentially, replicated the finding of O'Neill and many other studies have observed a remarkable proximity between the observed behavior and that predicted by the minimax solution (cf. Camerer 2011).

And so, what was observed as irrational behavior in early gaming experiments may, in fact, be an unintended consequence of a bad research design. Clearly, since conclusions of the early and the modern experiments are inconsistent, indicators of "rationality" used in these experiments cannot all measure what they were intended to measure. Fundamentally, then, the problem has to lie with the validity of a design.

We can only repeat after Aumann that the key to the problem of validity resides in "the motivation of the subjects and their understanding of the situation." Below we identify and comment on two principles that make this general observation a bit more specific.

Two Principles of Research Design

Consider, first, the following example. Say, for instance, that we test the level of proficiency in simple arithmetic and one of the problems on the test asks a subject to fill the blank in the expression "7 + 3 = …". Can we consider the answer to this question to be a valid indicator of arithmetic skills? Not necessarily. For the measure to be valid

[5] Note that the statistics reported here are averages over 25 pairs of subjects. Statistics for each pair of players were, obviously, different and some of them closer to the theoretically predicted than others. The variance across pairs was not large, however. For details see O'Neill (1987). Also note that assuming that subjects can solve for the minimax strategy and then play it would not be reasonable. But whether a subject can learn to play the minimax strategy is a different question. O'Neill's results suggest that he can.

it would have to be the case that all subjects understand signs "7" and "3" as denoting numbers "seven" and "three," that they understand operation "+" as denoting the arithmetical operation of addition, and that they all use the same system of arithmetic. These conditions need not be met—and the problem may not be purely academic. For example, most Americans write "seven" in a way that resembles how some Europeans write "one". Confusing the two is quite easy.[6] The moral is clear and we propose to tag it as *the first principle of research design* or *the Understanding Principle*: Everyone involved, both the subjects and the researchers, must have the same understanding of all concepts used by an indicator. This, of course, can only happen when such concepts are well defined. To put it differently, a shared common connotation of a concept is only possible if this concept comes from a properly constructed theory. The concept of a positive integer, a line or a square can be understood in a correct, hence identical way by different people because all these concepts come from properly constructed theories, arithmetic and Euclidean geometry in our case. Indeed, when the Understanding Principle is satisfied a test in arithmetic would get us as valid of a measure, as any measures we could ever get—with one proviso. This proviso relates to the subjects' incentives to reveal what we want them to reveal.

For a test to be a valid indicator of arithmetic skills, it is necessary to make sure that subjects have incentives to reveal what they think to be the proper answer to the question asked. In principle, they may not have such incentives. Consider, for instance, the following, a bit stylized, version of an actual classroom experience. Suppose that you teach two identical sections of a course. Suppose, moreover, that you give the same test in both sections. In the first section, however, you tell the students that test grades will count towards their class grade, while in the second section you say that you will use the test to see how useful it might be to review the material. Suppose you find out that the first section did very well on the test while the second one did very poorly. It is possible that the difference in the performance has nothing to do with a difference in the two groups' abilities; the difference may be caused by the different incentives the two sections faced. Some students in the second section could have deliberately given wrong answers to help you decide in favor of the review session. Had this been the case, the test would not have been a valid indicator of what they really knew. Clearly, incentives are crucial for the validity of a measure. Hence, the *second principle of research design*, or the *Incentives Principle*, must require that the design induces sincere answers.

Realizing potential problems with the seemingly perfect design of the "7 + 3 = ?" question should help us realize the scope of the validity issue. If we can have problems with this measure, then most standard measures used in social science research look hopeless in comparison. Consider, for instance, the case of measuring ideology. Can we safely assume that people asked about their ideological position on a left-right scale share a common understanding of the concepts involved in the question and have an incentive to reveal their true positions? Clearly, not. It is rather the opposite that strikes as a much more obvious alternative. What seems obvious, in other words,

[6] We have often seen an "American seven" being read as "one" (by a European) and "European one" being read as "seven" (by an American).

is that different people have different things in mind when thinking about political left and right and that they may not have an incentive to reveal their true position no matter what they perceive as left and right.

Neither identifying the two principles nor observing that the standard social science fails to satisfy them is particularly revealing. The two problems are far too obvious to surprise anyone. What is perhaps less clear is the type of research design that might overcome these problems.

A Classification of Research Designs with Respect to the Problems of Understanding and Incentives[7]

From our discussion above it should be clear that the two principles, and the two problems behind them, are independent. Given their independence, it is much easier to think about them separately, which probably explains the way solutions that solve these problems have been developed. Before we turn to the main topic of this paper—which is a design that solves the two problems at the same time—it will be instructive to look at research designs that solve them separately.

A useful way to think about the two principles of research design is by looking at their cross-classification. Given that the two principles are independent Table 2 provides a way to isolate four classes of designs.

Table 2

A Classification of Research Designs

The Understanding Principle \ The Incentives Principle	Satisfied	Not Satisfied
Satisfied	— proper connotation — no incentives to misrepresent	— proper connotation — incentives to misrepresent
Not Satisfied	— improper connotation — no incentives to misrepresent	— improper connotation — incentives to misrepresent

Note: "Improper connotation" means that the concepts do not have a sharp or a common connotation. "Incentives to misrepresent" means that the subjects have incentives to misrepresent their attitudes.

Table 2 provides a map to navigate the existing research designs in relation to the problem that is the subject of this paper. We began our discussion with a comment on the lower-right cell of Table 3 which represents studies that fail to satisfy both principles of research design. Unfortunately, they are very common in the social sciences. An increasing number of studies can be, however, placed in the upper right and the lower left cells of the Table 2. These studies satisfy one of the principles but not the other.

[7] The next two sections of the paper are the result of comments made by two anonymous referees. We are grateful for their most thoughtful suggestions.

Table 3

Examples of research designs that typically fall in the classification of Table 2

The Understanding Principle \ The Incentives Principle	Satisfied	Not Satisfied
Satisfied	?	Many studies in experimental game theory
Not Satisfied	Implicit indicators	Many standard designs like that of the National Election Study

We begin with a comment on studies that use properly defined indices some of which will end up in the upper right cell of Table 3, i.e., that fail to satisfy the Incentives Principle.

Experiments with gaming as well as many other experimental studies in economics have one common property that distinguishes them from many other studies in the social sciences—they use properly defined concepts. Just like it makes sense to assume that when asked about their age people will give an honest answer, it also makes sense to assume that when deciding between two actions they will pick an action—in the absence of other incentives—that gives them more money. The only condition we need to make sure of is that the subjects clearly see which of the two actions pays more.

O'Neill's study and its importance for experimental economics follows from the author's insightful conjecture that subjects' understanding cannot be assumed in a willy-nilly fashion, as in the early experiments, but has to be a part of a careful design. Using concepts with proper connotation is necessary but it is not sufficient to ensure that everyone understands them in the same way. In general it may be the case, just like in our example with number seven written in the American way, that different people will understand it differently and some, perhaps, won't understand it at all. But once we make sure that the rules of the game are properly understood by everyone, experiments like that become akin to tests in arithmetic—they can be given to anyone and meaningfully compared across individuals, and across cultures. This may well be the principal advantage of experimental economics that accounts for the exponential growth of this field and its ever increasing influence (Nobel Prizes in experimental economics have been awarded to Vernon Smith and Daniel Kahneman in 2002, and Alvin Roth in 2012.)

Experiments like that of O'Neill do not violate the Incentives Principle simply because the subjects' task in the experiment doesn't create any normative dilemma that may give them a reason to be insincere. This is true about great many other experiments on mixed-strategy equilibria (e.g., Binmore, Swierzbinski and Proulx 2001), like the O'Neill's study, on coordination games (e.g., Cooper, et al. 1990; Van Huyck, Battalio and Beil 1990), on learning (Stahl 2000) and other games in which players have no reason to be insincere.What happens if this is not true is a different

problem. Below we discuss several examples of studies that deal explicitly with the Principles of Understanding and Incentives.

Examples of Research Designs Addresssing the Problems of Incentives and Understanding

Dealing with incentives in surveys: "Respondents in surveys seem to lie for pretty much the same reasons they lie in everyday life—to avoid embarrassment or possible repercussions from disclosing sensitive information (…)" (Tourangeau and Tan 2007: 878). Indeed, asking some questions makes little, if any, sense. For instance, reviewing studies that compared self-reported and actual drug use Tourangeau and Tan (2007) report that 30%–70% of those whose post-survey urine tests proved using illegal drugs, reported that they did not use any. Naturally, it is easy to understand why people would lie about drug use. When it comes to using drugs, however, we don't really need to rely on subjects' answers—a simple urine test provides a perfectly valid assessment. The same cannot be said about assessing attitudes.

The problem with attitudes is that we do not have any objective testing method that would correspond to the urine analysis test. The weakness of studies in which we ask people to identify their attitudes is in subjects' conflating a true attitude with a sincere but inflated view of oneself with deliberate lie to make a favorable impression (Tourangeau and Tan 2007). In psychology, where careful experimentation is more common than in other social sciences, this is a well-recognized fact. Questions that may evoke insincere answers are referred to as sensitive questions. Psychologists attempted to solve the problem of sensitive questions earlier and their efforts were more extensive than those of other social scientists. The discussion that follows provides such a snapshot of this research.

One ingenious method of improving accuracy of self-reporting with sensitive questions was proposed by Warner (1965). His method works as follows: Imagine a random mechanism generating one of two possible answers e.g., "used drugs" and "did not use drugs." Consider, for instance, a roulette wheel with red and black pockets only (no green ones) where red corresponds to the first answer and black to the second. The subject is asked to spin the wheel and observe the outcome. All that is done in complete privacy. The only question the subject is asked to answer is whether he does or does not belong to the group identified by the random mechanism. Thus, if the outcome is "red," the question becomes "did you use drugs?" and a drug user answers "yes" while non-user answers "no." If the outcome is "black," the subject should invert his answers. A statistical estimation method is then used to determine the percentage of drug users.

Such randomized response methods have later been refined in a number of clever ways. One technique (Droitcour, et al. 1991) presents two groups of respondents with two lists of activities. One list consists of perfectly innocuous activities (e.g., went shopping, bought a new car, etc.), the other list adds to the first list a sensitive question (e.g., used drugs) which is the target of the study. In each group the subjects are merely

asked in how many of the listed activities they have engaged, say, in the last month; they don't have to identify which specific ones they were. Again, a straightforward estimation method can then be used to approximate the frequency of drug users.

Another common way psychologists use to study attitudes involves the so called implicit measures. As opposed to explicit measures, that reveal what is assessed, implicit measures are designed to conceal the assessed property and hence induce sincere responses (Petty, Fazio and Brinol 2009). The use of implicit measures has its roots in 1980s. According to Gawronski and De Houver (2012), the most frequently used paradigm is that of Greenwald, McGhee and Schwartz's (1998) implicit association test (or some version of it). The test, roughly, works as follows. To assess racial attitudes, for instance, subjects would be presented with positive and negative words and pictures of Black and White faces in all combinations (i.e., positive-Black, positive-White, negative-Black, negative-White) and would be asked to produce instantaneous associations. The basic idea is that quick and accurate responses will reveal subjects' true attitudes. Studies that followed the pathbreaking design of Greenwald, McGhee and Schwartz offer some ingenious refinements of the original design and provide many validity assessments of different designs. Gawronski and De Houver (2012) offer an excellenct review of the use of implicit measures in psychology.

Despite the widespread recognition of the problem of sensitive questions, the standard way to deal with the problem is to pretend that it does not exist. For a good measure of this attitude it suffices to consult any guides to constructing surveys. Cookbooks of survey methodology like *Improving Survey Questions: Design and Evaluation* (Fowler Jr 1995) or *Asking Questions: A Practical Guide to Questionnaire Design* (Sudman and Bradburn 1982), for instance, advise a "proper" wording of sensitive questions that would encourage respondent to reveal the truth. More specifically, to go with the running example of using drugs, the authors suggest to phrase a survey question as follows: "Using drugs is quite common among people of your age. Have you used any drugs in the last year?" We will leave this recommendation without a comment. The hopelessness of this advice reflects, perhaps, on the state of survey methodology in general.

The above methods, as much as they may be successful in solving the problem of incentive, do not, however, deal with the problem of understanding. A different, and unrelated, class of studies does precisely the opposite: it solves the problem of understanding but it does not address the problem of incentives. A good example of such research is experimental game theory.

Incentives problems in experimental game theory: Given that gaming experiments use properly defined concepts and behavior in games can reveal attitudes it would be natural to expect that experimental economics would take on this direction of research. It did. One of the most studied games is the so called ultimatum game. The simplest form of this game requires one player to propose a division of a fixed amount of money between herself and the other player who, in turn, can either reject or accept the offer. If the offer is accepted then the money is split as proposed, if it is rejected both players get nothing. Clearly, the second should accept any offer that gives him more than zero and therefore the first should offer him the minimal nonzero amount

possible. It does not happen, of course (c.f., Güth, Schmittberger and Schwarze 1982) and the reason is the players' concern of "fairness." The ultimatum game has many different versions[8] and much is known about factors affecting subjects' behavior. There is one aspect of this behavior that cannot be fully controlled however. Subjects in these experiments can clearly see that their actions reveal something about their norms and values and this means that they may not have sufficient motivation to reveal their true attitude. In fact, there is strong evidence showing that in studies in which subjects are guaranteed anonymity their "concern for fairness" drops dramatically (Hoffman, et al. 1994; Hoffman, McCabe and Smith 1998). In fact, just letting subjects know that their choices are no longer monitored by experimenters has the same effect on their behavior (Mironova and Whitt 2012).

As long as people can clearly see that their choices reveal something good or bad about them they may not want to reveal their true attitudes, no matter whether they remain anonymous or not. Even if we are perfectly anonymous lying to ourselves is a sufficiently strong incentive—especially that often we don't realize that we are lying at all. How we can go about solving this problem is a considerably more difficult question.

Problems with incentives and understanding: Measuring ideological positions or comparing the incomparable? As Inglehart and Klingemann (1979: 205) have noted "The term 'ideology' has caused much discomfort among social scientists. … [T]he history, multiple meanings and operational shortcomings of the concept have been discussed over an extended period of time." These problems are not specific to the concept of ideology. These are generic and standard problems with almost all normative concepts in the social sciences. The reason why ideology seems to have caused more discomfort than other concepts is because of its fundamental importance for political science and for sociology, and not because it is in any way qualitatively different from other normative concepts of importance for social sciences.

The Michigan school proceeds under an assumption that it makes sense to measure ideological positions directly. In the 2012 National Election Study, for instance, subjects were asked the following, very typical, question: *We hear a lot of talk these days about liberals and conservatives. Here is a seven-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. Where would you place yourself on this scale, or haven't you thought much about this?* (The same question was asked in earlier versions of the National Election Studies. They date back to 1977.)

It does not take much insight to doubt whether it makes sense to ask questions like that, and what, if anything, can be inferred from the answers. The reasons are clear. For instance, subjects may not understand the meaning of the concepts used in the question, or worse yet, they may think they understand the concepts while, in fact, they do not. Consequently, different people may have different things in mind when they think about liberal and conservative or left and right. In the most extreme

[8] A thorough review and summary of studies using different versions of ultimatum game are given by Camerer (2011).

case, people may have nothing in mind or have something that is completely opposite to what we assume they do. "Respondents who recognized the left-right dimension were then asked what they understood by 'left' and 'right' in politics. A sizable proportion of those respondents either *could not give any meaning* of the terms *or* else *completely* reversed their *meaning*" (Klingemann 1979: 230). The proportion was 20 percent in the Netherlands, 19 percent in Britain, 15 percent in the US, etc. There is no reason to expect this result would be significantly different at any point in time.

Even if two people have a precise understanding of what the left and the right means, it is close to impossible to expect that they both mean the same thing. It is enough to consider what "left" and "right" would mean to a Swede and to an American. But even if we assume that different people use different scales, which is to be expected, their scales may not be in any sense regularly related, for instance, linearly related, as is sometimes assumed for technical reasons as in e.g., Hinich and Munger (1994)[9]. And, as we move from comparing two individuals who come from a relatively homogeneous political culture to comparing cases from, say, Chile, Russia, Sweden and US, the problem becomes obviously unsolvable. And yet, it is not the most difficult problem we face when we ask people about their ideology.

The more difficult problem with measuring ideology may be related to the Incentives Principle of research design—subjects' motivation to reveal their true attitudes. Ideology, like many concepts in the social science, is a normative concept that involves value judgments. Values, however, are notoriously difficult to study. In many societies people are very reserved about values and norms when asked about them they may not want to reveal their true attitude but one that they think is expected of them. Indeed, incentives not to reveal true information may sometimes be enormous. Anyone familiar with totalitarian regimes will see the problem of the validity of public opinion surveys administered in times when a deviation from the official ideology could have implied a jail or a death sentence. But the problem of motivation goes deeper than that. People may also unknowingly misrepresent their attitudes. With rationalization and similar psychological mechanisms at work, it is easy to believe that what is proper is also true even if our answer may not possibly have any implications.

Consider, for instance, the following rephrased version of the National Election Study ideology question: *We hear a lot of talk these days about honest people and liars. Here is a seven-point scale from extremely honest to extremely dishonest. Where would you place yourself on this scale?* How much validity would you ascribe to this indicator of honesty? Clearly, very little. It seems outright silly to measure honesty by asking people how honest they are. But the general problem is common to the study of all normative concepts. Whenever values are involved, the validity of answers can never

[9] An acquaintance of ours has once argued that Milton Friedman could have been a better economist if he were not so possessed by the ideology of the left. He went on to explain in detail why Friedman and Lenin are obvious equivalents. His beliefs were not based on an inaccurate understanding of reality—they were just a function of his very stringent ideological scale. Relating his scale to that of some other people would most certainly require a nonlinear transformation.

be taken for granted. The design has to induce the revelation of true preferences or attitudes. In experiments with gaming this is induced through payoffs. However, when we want to measure honesty or political attitudes, fiscal rewards may not be sufficient to overcome the problem of incentives. When basic values and norms are at stake the revelation mechanism has to be different.

In any specific case, like measuring ideological positions, the solution to the problem of understanding and the problem of motivation may not be simple. It seems, though, that two constructive, general recommendations can be given. They are the subject of our next section.

The Role of Formal Theory in Solving the Problems of Incentives and Understanding

"If everyone had an ideology but no two persons had the same one, social scientists could make little use of ideology as an applied analytical concept. We could deal with ideology only as economists traditionally have dealt with consumers' tastes: admit they are profoundly important yet, lacking the capacity to measure or directly observe them, assume that they are constant and therefore can be ignored in causal analysis of changes in consumer behavior. Such an approach (more precisely, nonapproach) is unacceptable in dealing with ideology" (Higgs 1987: 45).

So, how do we solve the problems inherent in measuring ideology? Could it be that indicators used in the National Election Study, as bad as they clearly are, are the best that we can come up with? Or, is there a better solution?

We think that there is a way to construct a valid measure of ideological positions. The answer to how to do it is imbedded in part in the two examples above: O'Neill's study and the case of an arithmetical test. One reason both designs are capable of providing valid indicators is that both use only well-defined concepts. This, by definition, induces a clear and a commonly shared understanding of the problem among the subjects. But providing sharp definitions of all concepts involved in the research problem is the very objective of what in political science and sociology is often referred to as a formal theory. (Economics and sciences exhibit a better sense by skipping the predicate "formal" or, equivalently, not using the name "theory" on constructs that bear this name elsewhere.) Hence, the connection between the concept of a formal theory and a prescription for a valid indicator.

Practical advice for the construction of a valid indicator would be to use concepts that either have a sharp and a commonly recognized denotation (like numbers, letters, proper names, etc.) or concepts that acquire such denotation through a set of instructions given to the subjects. In other words, we propose to think about an indicator as a version of a formal theory, one that is sufficiently straightforward to be easily comprehensible, yet sufficiently detailed to give all concepts a sharp denotation. An indicator like that would, by definition, solve the first problem of research design, the problem of common understanding. This is our first advice, one that applies to the problem of consistency.

A possible solution to the second problem, the problem of motivation, is less clear. One aspect of the problem is how to induce someone to reveal an ideological position when no incentives are strong enough. Another aspect is how to uncover a true attitude when subjects' own psychological mechanisms make them believe in a false one. To solve the two problems, we need a design that would make a subject unaware of the normative significance of his answers/choices. If we were to think about an indicator as a form of a formal theory, deductions leading from subjects' choices to their normative interpretation would have to be sufficiently complex in order to be practically untraceable by non-experts. In other words, we would not want the subjects to be able to link their answers to the problem posed in the survey with the answer's normative meaning. This, essentially, is the content of our second advice, one that applies to the problem of motivation. How the two specific advices of this section can be implemented is the subject of our next section.

A Research Design and a Theory Behind It

Consider a situation that can be easily, yet in a very precise way, described by the following instruction.

The jury in a competition consists of three jurors who have to select three candidates, A, B and C, for three main prizes. Naturally, verdicts of different jurors can be different, perhaps even quite discordant. Each verdict of a juror is written down as follows: the first letter identifies the candidate, proposed by this juror for the first prize, the second letter, for the second prize, and so on. For example, ABC denotes a verdict to award the first prize to candidate A, the second, to candidate B, and the third, to candidate C.[10]

Suppose now that the three jurors in the competition return the following verdicts: ABC, ABC, BCA. (Let's label this particular situation of individual preferences on three alternatives as Profile IV. The reason for the unorthodox numbering will be revealed later.) Taking into account opinions of all three jurors, what would you consider to be the optimal verdict of the jury?

We think it reasonable to assume that it does not take any scholarly knowledge to understand the instructions, much like in the case of O'Neill's card game design. Of course, readers familiar with the discipline of public choice would recognize, in what was just described, the standard problem of the field: how to aggregate individual preferences of citizens into the single best outcome for the society. The famous answer to this problem, known as Arrow's Theorem, says that there is no satisfactory solution; any rule of aggregation has to violate some conditions that are fundamentally important from the normative point of view.

For the practice of social choice, the inherent lack of desirable solutions is, of course, undesirable. Yet the absence of good solutions may be quite useful for other purposes. For instance, if no aggregation method is ethically neutral or value free, by

[10] This is an initial part of an instruction used in a study by Lissowski in 1988. The study and its results are described in Lissowski and Swistak (1995).

choosing solutions to aggregation problems people may be revealing values that lead them to these choices. And if this is so, the social choice setup, like the one described above, may be used to uncover these values.

But let's return for a moment to the case of Profile IV where {ABC, ABC, BCA} is the set of individual preferences. What would *you* consider to be the best social outcome? When this question was asked in Japan 36% of subjects prescribed ABC and 63% picked BAC. In a sample from Poland the distribution of choices was drastically different: 86% chose ABC and only 11% chose BAC.[11] It clearly looks like an average choice in the two countries was driven by different values. But what are the values that can be inferred from the different social outcomes? Can you tell, for instance, what sort of a political attitude underlies the choice of ABC and how it differs from the attitude that leads to the choice of BAC? If you cannot, then this is precisely what is needed in a research design to overcome the problem of a subject's motivation. Recall that the necessary property of a design that deals with normative issues should be the lack of any obvious connection between the response of a subject and the normative meaning of her response. Indeed, we believe that using a design like that may well be the only way to avoid both conscious and subconscious distortions in subjects' responses.

A theory needed to link social outcomes with their normative meaning is somewhat complex and a bit technical. It was carefully laid out elsewhere (Lissowski and Swistak, 1995) and this is where we would like to refer the interested reader for all details. We believe, however, that the basic idea behind the construction is very intuitive and a less technical reader may well be satisfied with the following sketchy explanation.

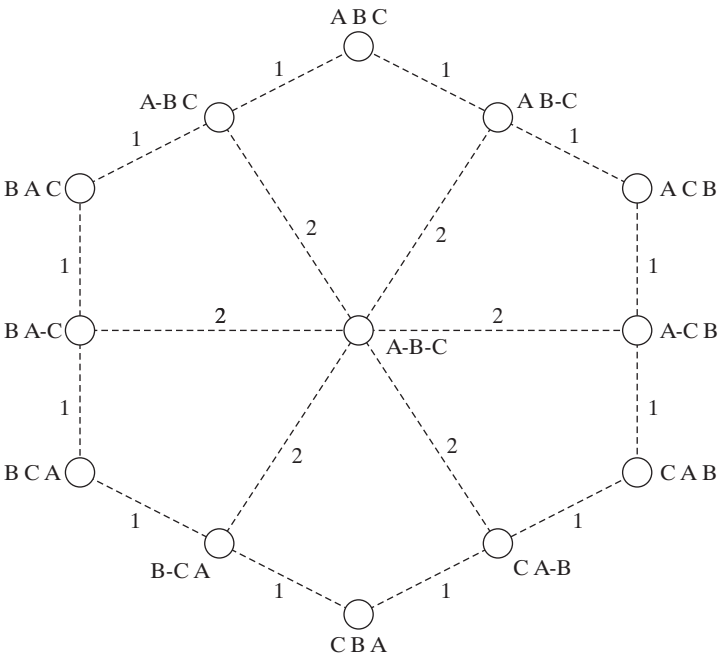
Consider again the set of individual preferences {ABC, ABC, BCA}. For this profile all standard social welfare functions prescribe ABC as the social outcome. However, an obvious ethical "problem" with the choice of ABC is that while ABC perfectly reflects the first two preferences, it completely ignores the third one. To account for the fact that one order "perfectly reflects" or "completely ignores" another one, we need a measure of distance on preferences. Interestingly, it can be proved (Lissowski and Swistak, 1995) that such measure is uniquely given by a set of normative conditions that describe the social choice setup used in the study (anonymity, neutrality, etc.).

Figure 1 illustrates this distance measure. It depicts a map with all thirteen theoretically possible social outcomes. If an order in a social outcome is not strict and alternatives A and B are tied, we write A-B. A-BC, for example, means that A is as good as B, and they are both better than C. The shortest distance in this map is taken as a unit and the distance between any two orderings is defined as the length of the shortest path that connects them. Thus, the distance between ABC and A-B-C, for instance, is 3 (1 + 2). This distance measure constitutes the first part of the theory that links social outcomes with values that underlie their choice. The path to the second part of this theory leads through the following, rather remarkable, empirical finding.

[11] Details of these studies and their results are given in Lissowski and Swistak (1995).

Figure 1

Distances Between all Rankings of Three Alternatives



Optimality of Choices

Designs that are similar to the one described above have been used in a few studies (Lissowski and Swistak 1995). The most extensive of these studies was the study by Lissowski conducted in 1988 on 144 undergraduate students in Warsaw. Results that are reported below come from this study.

In one part of the Lissowski study the subjects were asked to choose the best social outcome in eight different profiles.[12] Profile IV was one of the eight profiles used in the study. Each profile was identified by a different set of jurors' preferences. The instruction, fragments of which were quoted above, depicted three candidates judged by three jurors in a competition in which first, second, and third prices were to be awarded, however ex aequo prices were also possible.

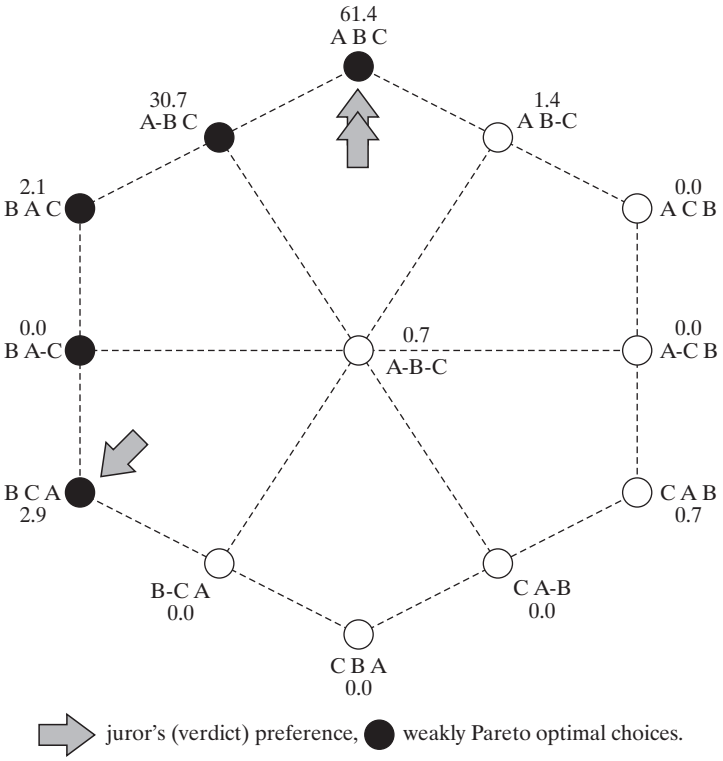
The first part of a theory that may help us explain the nature of choices made by the subjects in the study is the distance geometry of Figure 1. Using this geometry we can depict profiles together with the distribution of subjects' choices. Figure 2 shows this picture for the case of Profile IV. While Figure 2 has the same basic structure as the graph of distances in Figure 1 it also includes a number next to each order which stands for the percentage of subjects who chose this order as the best social

[12] If we restrict our attention to strict orderings of the alternatives by the jurors the set of eight profiles exhausts the set of all substantively different and nontrivial situations (i.e., the profile of identical preferences and the Condorcet profile were excluded). See Lissowski and Swistak (1995) for explanation.

outcome. In addition, jurors' preferences (ABC, ABC, BCA) are marked by arrows. The distribution of subjects' choices can thus be seen as imbedded in the geometry of distances among orders and the geometry of jurors' preferences.

Figure 2

Distribution of Choices (in %) for Profile IV



While it may take a moment to become familiar with the structure of Figure 2, once such familiarity is attained, a striking regularity comes through this somewhat convoluted structure—*almost all choices fall somewhere in between the two most distant arrows* (preferences of the jurors). In Profile IV the percent of these choices is 97.1. What normative property can explain the meaning of this geometric regularity?

Consider any social outcome in Profile IV and describe it by a vector of three numbers: distances from the three preferences of jurors (marked by the arrows). Given that jurors are abstract undifferentiated entities, hence impartiality is induced by the design, we can identify each social outcome by an ordered vector of distances with coordinates arranged from the smallest to the largest. Hence A-BC, for instance, would be identified by the vector (1,1,3) and ACB, by (2,2,6). Now, if we were to compare A-BC with ACB, it is easy to see by comparing (1,1,3) with (2,2,6) that, the first order is better in all three coordinates (the closer the distance the better the

choice). Formally we say that the first vector dominates the second one in the weak Pareto sense.[13]

Consider now all social outcomes that are weakly Pareto optimal. The geometric regularity of most choices "falling in between the two most distant arrows" corresponds, roughly speaking, to most choices being weakly Pareto optimal. Now that the normative condition of (weak) Pareto optimality is specified we can revisit the data and ask what percentage of subjects' choices were weakly Pareto optimal or, in short, optimal.

Recall first that there were 144 subjects in the study. The study concluded with 140 valid responses. Each of the 140 subjects chose one social outcome for each of the eight different profiles used in the study. Hence 1120 choices were recorded in total. Remarkably, only 52 of the 1120 choices *were* not optimal. In other words, 95.4% of all choices *were* optimal. This is a rather striking regularity, certainly one that you would not expect of "the truly unsophisticated cretins."

However the more important aspect of subjects' choices is the ideological meaning of the choices they made. This, in fact, is the main objective of the design, which still remains to be explored.

Ideological Consistency of Choices

How can we use the social outcomes chosen by the subjects to identify subjects' ideological positions? The following short reflection on the nature of politics and the role of ideology may help us see the connection between a choice and its ideological meaning.

The main purpose of politics is redistribution and the justification for any pattern of redistribution is provided by ideology. The two main ideologies are liberal and conservative or, rather, left and right. To see the essence of the two, think about a society as an ordered vector of interpersonally comparable utilities. Suppose that utilities are ordered from the lowest, these correspond to the worst off members of the society, to the highest. Roughly speaking, ideologies of the left profess maximizing utilities of the left end segment of this vector whereas ideologies of the right profess maximizing utilities of the right end. This simplified view of the problem, though likely objectionable for a political philosopher, turns out to be just the right type of a theoretical insight needed for interpreting subjects' choices in the study.

Recall that each social outcome in the social choice problem can be described by a vector of distances. Distances are, obviously, inversely proportional to the utilities of the social outcome for the three jurors. If we now translate these distances into

[13] In general, an ordered vector $(d_1,\ldots,d_n)$ of distances is better in the weak Pareto sense than $(d_1{}^*,\ldots,d_n{}^*)$ if and only if $d_i < d_i{}^*$, for all $i = 1,\ldots,n$. That is, a vector which is better in the weak Pareto sense has smaller distances in all coordinates. It is useful to consider a more stringent notion of optimality—the condition of strong Pareto optimality for ordered vectors. We will say that an ordered vector $(d_1 \ldots d_n)$ is better than $(d_1{}^*,\ldots,d_n{}^*)$ in the strong Pareto sense if and only if $d_i \leq d_i{}^*$, for all $i = 1,\ldots,n$ and for at least one j, $d_j < d_j{}^*$. A vector that is better in the strong Pareto sense has all distances smaller than or equal to, and at least one of them strictly smaller. We denote this relation by $\succ$ and call it a strong Pareto relation.

utilities the problem of social choice translates into the problem of distributive justice. What remains to be done, in this representation, is to identify categories of social outcomes that correspond to the ideologies of the left and the right. But, from what was said above, the task is easy: choices that maximize proper left end intervals of the utility vector should be identified as leftist and choices that maximize proper right end intervals as rightist.[14] This gives us a simple conceptual framework to assess the ideological nature of subjects' choices.

To judge subjects' ideological positions, however, we have to see if there is any consistency in their choices across the different profiles. Consistency of choices across situations and across time[15] is, after all, a necessary property of a political attitude. Hence someone who opts for a leftist solution in one situation and a rightist solution in another cannot possibly be seen as ideological, i.e., cannot be meaningfully referred to as homo politicus. The stage is set now to separate homo politicus from a "truly unsophisticated cretin."

To judge ideological consistency of choices we will restrict our attention to the set of subjects whose choices can be explained by the leftist and the rightist categories in all eight profiles (80.7% of the sample). It turns out that in this set only 6.2% of subjects were inconsistent—the stunning 93.8 percent were *consistent*.[16] The magnitude of the regularity is so remarkable that it can rival, we believe, the amazing result of the Campbell et al. (1960) study where only about 2.5% of the electorate showed an evidence of clear ideological thinking. The difference, of course, is that the results of the two studies point in the opposite directions: One result suggests perfect absence of ideological thinking, the other, its clear presence. It is difficult to miss the analogy between these two studies and the previously discussed experiments on equilibria in zero-sum games where O'Neill's finding that subjects perfectly converge on the equilibrium strategy was in direct opposition to the earlier studies. It appears that in both cases the key to the inconsistent conclusions lies with the proper research design. Once the subjects know what they are asked about and have no incentives to lie the irrational cretins turn to rational prodigies. The joke, it seems, is on us.

And so, we conclude that the seemingly elusive homo politicus may exist after all. We just need a proper design to uncover his existence. The two key elements of the design involve two properties of a formal/scientific theory: properly defined

[14] Lissowski and Swistak (1995) refer to the two classes of solutions as generalized Rawlsian and generalized conservative solutions. Formally the two classes are defined as follows. Take the following relation $\succ_{Rk}$ defined on ordered vectors of distances, for any k = 1,…,n: $(d_1,\ldots,d_n) \succ_{Rk} (d_1{}^*,\ldots,d_n{}^*)$ $\underline{\text{iff}}$ $(d_{n-k+1},\ldots,d_n) \succ (d_{n-k+1}{}^*,\ldots,d_n{}^*)$. The $\succ_{Rk}$ relation defines the generalized k-th degree Rawlsian relation of justice. Denote the set of optimal elements of $\succ_{Rk}$ as Rk. The set of generalized Rawlsian solutions is defined as the sum: $R1 \cup R2 \cup \ldots \cup R(n-1)$. The definitions are analogous for the conservative set of solutions. The only difference, of course, is that the corresponding generalized k-th degree conservative relation of justice is defined as $\succ_{Ck}$, for k = 1,…,n: $(d_1,\ldots,d_n) \succ_{Ck} (d_1{}^*,\ldots,d_n{}^*)$ $\underline{\text{iff}}$ $(d_1,\ldots,d_k) \succ (d_1{}^*,\ldots,d_k{}^*)$.

[15] We do not want to imply that a change of values, norms, or attitudes is not possible. Allowing for a sufficiently long time interval, a change is, obviously, possible. Our position here is that if a change is observed over a very short time period (honest today, dishonest tomorrow, for instance) then it indicates the lack of a value (norm, attitude) rather than a change from one value to another.

[16] If subjects whose choices can be explained by the leftist and the rightist categories in all eight profiles were to pick the social outcome at random, only 7.6% of them would have been consistent across all eight profiles.

concepts and the presence of difficult deductions. The first property implies that subjects understand what they are asked about, the second property prevents them from seeing the link between their answers and the attitude these answers reveal.

We believe that when studying normative issues, like ideological attitudes, the lack of a comprehensible link between the answer and its normative meaning is the necessary property of a valid design. The absence of this link makes it impossible for a subject to distort, consciously or subconsciously, his true attitudes.

Conclusion

The lesson from the studies described in this paper is that the answer to the problem of a valid research design is to be found, ultimately, in the proper form of practicing science—a formal theory, as some political scientists and sociologists call it.

Some empirical findings on which we base the argument may be thought of as partial and preliminary but they show exceptionally strong regularities that warrant future research.

Studies that we have used to present the argument provide, in our opinion, the clearest examples of the two principles of research design: the Understanding Principle and the Incentives Principle. These studies stand in a sharp contrast to the standard studies that refuse to deal with the two problems of research design.

It is our belief that a solution to the problems of incentives and understanding leads through the proper form of a scientific theory. Just as a proper theory construction carries the answer to many pitfalls of theoretical research, it may also prove to be the key solution to notorious problems of survey and experimental research in the social sciences.

Our preliminary analysis suggests that the image of the homo politicus and homo sociologicus that will emerge from the properly designed research may be in stark contrast to what we came to believe is true.

References

Aumann, Robert. 1987. Game Theory, in: *The New Palgrave: A Dictionary of Economics*. London: Macmillan.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Binmore, Ken, Swierzbinski, Joe and Proulx, Chris. 2001. Does Minimax Work? An Experimental Study. *The Economic Journal*, 2001: 445–464.

Camerer, Colin F. 2011. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton NJ: Princeton University Press.

Campbell, Angus, Converse, Philip E., Miller, Warren E. and Stokes, Donald E. 1960. *The American Voter*. New York: Wiley.

Converse, Philip E. 1964. The Nature of Belief Systems in Mass Public, in: David E. Apter (ed.), *Ideology and Discontent*. New York: Free Press.

Cooper, Russell W., DeJong, Douglas V., Forsythe, Robert and Ross, Thomas W. 1990. Selection Criteria in Coordination Games: Some Experimental Results, *The American Economic Review*, 1990: 218–233.

Davis, Douglas D., and Holt, Charles A. 1993. *Experimental Economics*. Princeton: Princeton University Press.

Droitcour, Judith, Caspar, Rachel A., Hubbard, Michael L., Parsley, Teresa L., Visscher, Wendy and Ezzati, Trena M. 1991. The Item Count Technique as a Method of Indirect Questioning: A Review of Its Development and a Case Study Application, in: Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz and Seymour Sudman (eds.), *Measurement Errors In Surveys*. New York: Wiley.

Enelow, James M., and Hinich, Melvin J. 1994. *The Spatial Theory of Voting*. Cambridge: Cambridge University Press.

Fowler Jr, Floyd J. 1995. *Improving Survey Questions: Design and Evaluation*. Thousand Oaks CA: Sage Publications.

Gawronski, Bertram, and De Houver, Jan. 2012. Implicit Measures in Social and Personality Psychology, in: Harry T. Reiss and Charles M. Judd. *Handbook of Research Methods in Social and Personality Psychology (2nd edition)*. New York: Cambridge University Press.

Greenwald, Anthony G., McGhee, Debbie E. and Schwartz, Jordan L. K. 1998. Measuring Individual Differences in Implicit Cognition: the Implicit Association Test, *Journal of Personality and Social Psychology*, 1998: 1464–1480.

Güth, Werner, Schmittberger, Rolf and Schwarze, Bernd. 1982. An Experimental Analysis of Ultimatum Bargaining, *Journal of Economic Behavior & Organization*, 1982: 367–388.

Higgs, Robert. 1987. *Crisis and Leviathan*. Oxford: Oxford University Press.

Hinich, Melvin J., and Munger, Michael C. 1994. *Ideology and the Theory of Political Choice*. Ann Arbor: The University of Michigan Press.

Hoffman, Elizabeth, McCabe, Kevin and Smith, Vernon. 1998. Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology, *Economic Inquiry*, 1998: 335–351.

Hoffman, Elizabeth, McCabe, Kevin, Shachat, Keith and Smith, Vernon. 1994. Preferences, Property Rights, and Anonymity in Bargaining Games. *Games and Economic Behavior*, 1994: 346–380.

Ingelhart, Ronald, and Klingemann, Hans-Dieter. 1979. Ideological Conceptualization and Value Priorities, in: Samuel H. Barnes, Max Kaase and Klaus R. Allerbeck (eds.), *Political Action*. Beverly Hills: Sage.

Inglehart, Ronald. 1990. *Culture Shift in Advanced Industrial Society*. Princeton: Princeton University Press.

Kahneman, Daniel, Slovic, Paul and Tversky, Amos. 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kaminski, Marek M. 1999. How Communism Could Have Been Saved. Formal Analysis of Electoral Bargaining in Poland in 1989. *Public Choice*, 1999: 83–109.

Klingemann, Hans-Dieter. Measuring Ideological Conceptualizations, in: Samuel H. Barnes, Max Kaase and Klaus R. Allerbeck (eds.), *Political Action*. Beverly Hills: Sage.

Lissowski, Grzegorz, and Swistak, Piotr. 1995. Choosing the Best Social Order: New Principles of Justice and Normative Dimensions of Choice. *American Political Science Review*, 1995: 74–96.

Mironova, Vera, and Whitt. Sam 2012. *International Peacekeeping, and Positive Peace: Experimental and Survey Evidence from Kosovo*.

Niemi, Richard D. and Weisberg, Herbert F. 1984. *Controversies in Voting Behavior*. Washington D.C.: CQ Press.

O'Neill, Barry. 1987. Nonmetric Test of the Minimax Theory of Two-Person Zerosum Games. *Proceedings of the National Academy of Sciences*, 1987: 2106–2109.

Petty, Richard E., Fazio, Russell H. and Brinol, Pablo. 2009. The New Implicit Measures, in: Richard E. Petty, Russell H. Fazio and Pablo Brinol (eds.), *Attitudes: Insights from the New Implicit Measures*, 3–18. New York: Psychology Press, 2009.

Piattelli-Palmarini, Massimo. 1994. *Inevitable Illusions*. New York: Wiley.

Rapoport, Amnon, and Boebel, Richard B. 1992. Mixed Strategies in Strictly Competitive Games: A Further Test of the Minimax Hypothesis, *Games and Economic Behavior*, 1992: 261–283.

Rapoport, Anatol, Guyer, Melvin J. and Gordon, David G. 1976. *The 2 × 2 Game*. Ann Arbor: The University of Michigan Press.

Stahl, Dale O. 2000. Rule Learning in Symmetric Normal-form Games: Theory and Evidence. *Games and Economic Behavior*, 2000: 105–138.

Sudman, Seymour, and Bradburn, Norman M. 1982. *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.

T o u r a n g e a u, Roger and T a n, Ting. 2007. Sensitive Questions in Surveys . *Psychological Bulletin*, 2007: 859–883.

V a n H u y c k, John B., B a t t a l i o, Raymond C. and O B e i l, Richard. 1990. Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure. *The American Economic Review*, 1990: 234–248.

W a r n e r, Stanley L. 1965. Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, 1965: 63–69.

Biographical Notes:

Marek Kaminski is Associate Professor of Political Science, Affiliated Professor of Economics and member of the Institute for Mathematical Behavioral Sciences at the University of California. Before moving to California in 2002, he was Assistant Professor at New York University. He holds graduate degrees in sociology, applied mathematics and political science from University of Warsaw and the University of Maryland at College Park. His research interests include democratization and political consequences of electoral laws, models of prison subculture and axiomatic models of allocation of scarce resources. His publications include *Games Prisoners Play* (Princeton 2004, Polish edition *Gry Więzienne*, Oficyna Naukowa 2006).

E-mail: marek.kaminski@uci.edu

Grzegorz Lissowski is Professor of Sociology, Head of Chair of Statistics, Demography, and Mathematical Sociology, the Institute of Sociology, University of Warsaw, Poland. Main research topics concern mathematical sociology, decision theory, social choice theory, and statistics. His most recent publication is the book *Principles of Distributive Justice*, Verlag Barbara Budrich 2013.

E-mail: g.lissowski@uw.edu.pl

Piotr Swistak is an Associate Professor in the Department of Government and Politics and in the Applied Mathematics, Statistics and Scientific Computation Program at the University of Maryland, College Park, Maryland, USA. He has published in journals in sociology, psychology, political science, philosophy and mathematics. His publications include: Jonathan Bendor and Piotr Swistak. *Types of Evolutionary Stability and the Problem of Cooperation, Proceedings of the National Academy of Sciences, U.S.A.*, 1995, Vol. 92, pp. 3596–3600; Grzegorz Lissowski and Piotr Swistak. Choosing the Best Social Order: New Principles of Justice and Normative Dimensions of Choice, *American Political Science Review*, 1995, Vol. 89, No. 1, pp. 74–96; Jonathan Bendor, Rod Kramer and Piotr Swistak. Cooperation under Uncertainty, *American Sociological Review*, 1996, Vol. 61, No. 2, pp. 333–338; Jonathan Bendor and Piotr Swistak. The Evolutionary Stability of Cooperation, *American Political Science Review*, 1997, Vol. 91, No. 2, pp. 290–307; Jonathan Bendor and Piotr Swistak. The Evolution of Norms, *American Journal of Sociology*, 2001, Vol. 106, No. 6: 1493–1545. All of his research concerns explaining human behavior.

E-mail: pswistak@umd.edu