Soc Choice Welfare (1987) 4:43-62

Social Choice and Welfare

An Interview with Kenneth J. Arrow

J. S. Kelly

Department of Economics, Syracuse University, Maxwell Hall, Syracuse, NY 13210, USA

The following is an edited transcript of an interview conducted on March 4, 1986 with Professor Arrow while he was visiting Syracuse University to deliver the Frank W. Abrams Lecture Series to be published as *The Uncertain Future and Present Action* by Syracuse University Press. This interview was to elaborate on his description, presented in Volume 1 of his *Collected Papers* (Harvard University Press, 1983), of the origins of his work in collective choice theory.

JK. You started off the story in the collected papers with remarks about studying relational logic while you were in Townsend-Harris High School in New York City.

KA. Not in high school in the sense of in my high school courses, but during this period I was an omnivorous reader and got into all sorts of things. One of them was Bertrand Russell's *Introduction to Mathematical Philosophy* and it made a tremendous impression on me. It was the idea of logic that was in there. I don't really recall, for example, if there was a formal definition of a relation as a set of ordered pairs, but I learned the ideas of mathematical logic and its applications to mathematics in Russell's book. It seems to me that I also read one or two other logic books around that time.

JK. Later, when you went to City College of New York as a mathematics major, you encountered more mathematical logic.

KA. Yes, but again the logic study was on my own, there were no courses in it. I don't really remember exactly what I read. I remember once taking out the *Principia Mathematica* but of course it's not the sort of thing one really reads from. I was looking up some theorems in it and things like that. I really am not prepared to tell you what I read, but at some point things like the idea of defining rational numbers by ordered pairs and equivalence classes by ordered pairs was something I got to know. I was fascinated by this and used to aggravate my professors by writing out proofs in very strictly logical form, avoiding words as much as possible and things of that kind.

JK. You did take a formal course with Tarski in the Philosophy Department; how did you happen to take that course?

KA. Yes. Well, I *knew* that Alfred Tarski was a great and famous logician and there he was in my last term in school and obviously I was going to take a course with Alfred Tarski. It turned out he had two courses. One was a kind of introductory course and I felt I knew more than *that*. The other course he gave was in the calculus of relations. To say it was in the *calculus* of relations meant that he gave an axiomatic treatment of relations, although he motivated it of course by motivating the axioms. You never had xRy; you only had R and S and T. You see, he never mentioned *individuals* in the formal theory. He had an axiomatic theory like an axiomatic treatment of set theory. Relations have some special aspects, in particular the idea of relative product, RS. If there is a z such that xRz and zSy, then $xRSy$. The relative square, $R^2 = RR$ is especially interesting; if the relative square is included in R you have transitivity.

So it was a fascinating thing, although it was really very elementary; really very easy. The concepts were not very subtle compared with the deep things he was working on like the truth principle.

JK. At this point you were involved in translating some of Tarski's work.

KA. He wrote a textbook called *Introduction to Logic* [1] which is one of the modern treatments, modern as of 1940. It had been published in German, may even have been orginally published in Polish. I didn't translate it. What happened was he had a translator and I read the proofs. I was just finishing college and he asked me to read the proofs for him. He didn't know any English, you see. This was the interesting thing. He came to this country in September, 1939 for some kind of congress or conference and was trapped here by the outbreak of the war. He knew Polish, he knew German, but he didn't know any English so he spent the Fall term learning some English so he could teach us in the Spring. At first we couldn't understand a word he was saying but after about a week or so we began to catch on and we realized it wasn't *his* rate of progress it was *our* rate of progress that was relevant. His stresses were all wrong. He was aware of this and therefore felt he couldn't proofread in English. It's rather interesting as a coincidence that the translator was a German philosopher named Olaf Helmer and Helmer comes back into my story eight years later.

It's interesting... Tarski, although his English was weak, had a very good sense of language and he kept on asking me "Is that really good English?" Not in the sense of being grammatically correct, but, well for example, Helmer was very fond of using the word "tantamount" and Tarski got the feeling that somehow it's not a word used very often. Actually his instincts for language were extremely good. I suppose that was connected with his general work on formalizations and metalanguages. Anyway, I was just a proofreader.

JK. You write that as a graduate student at Columbia you spent time as an exercise translating consumer theory in the logic of relations and orderings. What got you started on that and what did you get out of it?

KA. I went to Columbia because . . .well there were several problems. One was that we were extremely poor and the question of going anywhere depended on resources. Columbia had the great advantage, of course, that I could live at home, which wasn't true anywhere else. I didn't get any financial support for my first year, none at all.

But another of the things I had learned on my own at college was mathematical statistics and I really had become fascinated with it. There was a course in statistics [at CCNY]; the teacher, a man by the name of Robinson, had no *real* knowledge of it I would say, basically – I won't even say he had a good reading list – but he did list one book, J. F. Kenney [2] if I remember correctly, which happened to have an excellent bibliography. It was not one of those cookbooks in statistics but actually did have some attempts at mathematics. Kenney had references to R. A. Fisher and gave you enough to get you interested. So I started reading Fisher and one of the first things was trying to work out his derivation of the distribution of the correlation coefficient under the null hypothesis, which was an integration in n-dimensional space. In Fisher it was done by intuition. I mean it's rigorous if you're sufficiently sophisticated; to me it was gibberish. But I knew enough multivariate calculus to be able to translate it into rigorous form, at least a form that I understood, and then I could see that he really was right. But I couldn't see it the way he wrote it. Then I suppose because of my logical background what was really important was reading the Neyman-Pearson papers which were then new and written in rather obscure places, but they were available in the [CCNY] library. From Fisher alone, I think I would have been hopelessly confused about the logic of statistical tests, although Fisher was great on deriving distributions.

So, I knew I wanted to study mathematical statistics, which however was not a field, not a Department at Columbia. It was spread out in other Departments. I knew that Hotelling was one of the major figures, but he was in the Economics Department. I rather naively thought I would study mathematics and then would take the statistics from Hotelling. I had no interest in Economics.

I was in the Mathematics Department, taking courses like Functions of a Real Variable, but I was going to take courses from Hotelling. In the first term he happened to give a course in Mathematical Economics. So out of curiosity I took this and got completely transformed.

The course to an extent revolved around Hotelling's own papers. But, as it happens, they were kind of central. He gave a rigorous derivation of supply and demand. There was one paper on the theory of the firm, one on the theory of the consumer [3, 4]. And he gave a rigorous derivation of demand functions in the consumer theory paper and derived the Slutsky equations. I think he knew about Slutsky's work, though I'm not sure he actually referred to Slutsky. So, anyway, this was one of the best papers around at the time. It's now a staple of our literature but then really was novel. One of the things, he was a very, *very* strong ordinalist, emphasized that all these results were invariant under monotone transformations, which was not a normal practice in economics at that time. Of course, all those who were coming of age, like Paul Samuelson, would jump to that position; it was the normal position of the avant garde.

Well, the idea was that it was an *ordering*. It was clear that what they were saying was "x is better than y" and that this is a transitive relationship. And I recognized that there were certain continuity axioms that had to be added to that. I was already familiar with that because there were certain similar things in the foundations of probability theory. In fact I think I worked that out for myself. I was playing around once in college trying to work out an axiom system for probability theory, that was work on an Honors paper or something, and I ran across a set of axioms by Karl Popper. Research methods were pretty primitive; I looked through the Union catalog and there was a reference to an article [5] in *Mind* by Popper. I realized that his axiom system really couldn't explain certain things that we take for granted like the fact that cumulative distributions have a one-sided continuity property. So I realized that you need some kind of extra continuity axiom and I sort of invented countable additivity all by myself. Later, of course, I found that Kolmogoroff and others had done this, but I could see there had to be an axiom.

So I was kind of familiar from having worked it out there that you needed these continuity axioms in order to close your preference theory system. It was easy to provide and I suppose others were doing the same. I could also see that while it was clarifying for me, it was hardly a contribution to knowledge because all I was doing was translating to a language that I knew. At least it got me thinking; whenever I saw a U for a utility function I translated to a preference ordering.

In fact one thing that struck me as an interesting problem – this is digressing a bit, but not entirely – why should there be a utility function representing an ordering? Hotelling had never really asked that question. Although he emphasized that the indifference map was the primitive, and the utility function only represented it, he didn't really ask "Why should you have a representation in terms of numbers?" I was really thinking about this problem when I happened to run across some papers [6] by Herman Wold who gave what he called a "Synthesis" in some papers in *Skandinavisk Aktuarietidskrift* which gave a long treatment of demand analysis which did have essentially an axiomatic point of view. There he said you've got to prove there is a utility function representation. He was the first person I know to realize, in print, that this was a problem. He gave an answer, extremely weak because he needed strong assumptions.

Anyway, then I switched to Economics from Mathematics. I had gone to Hotelling asking for a letter of recommendation for a fellowship in the Mathematics Department and he said, "Well, I'm sure I dont't have any influence in the Mathematics Department, but if you should enroll in Economics, I've found in the past they are willing to give one of my students a fellowship." I was bought.

Incidentically, I impressed him on about the second day of the class because he was fascinated by Edgeworth's taxation paradox; in fact his paper on the theory of the firm was called "Edgeworth's Taxation Paradox and the Nature of Supply and Demand Theory" [3]. Consider a case where there are first class and third class railroad tickets as in the English system. It turns out that if you impose a tax on one ticket then, with suitable demand functions, you could lower the price of both commodities. At the time there was a lot of excitement about that; the public finance people were pooh-poohing it, saying, "How can this be?" It had to do with the nature of inter-related demand curves and that was the big thing Hotelling stressed, that demand functions depended on *n* variables, not one variable. But he said he was puzzled by the fact that he had never been able to produce an example of Edgeworth's paradox with linear demand functions. So I sat down and wrote out the conditions for linear demand functions to yield the paradox; these conditions were certain inequalities on the coefficients and the inequalities were inconsistent. So I came in the next day and showed it to him. Really it was just a few lines, but from that point on he was really impressed with me. It was an extremely easy calculation, but thinking in inequality terms was not common. Little pieces were quite easy to prove, but you couldn't do it in the mechanical fashion which you were doing with, say, solving simultaneous equations or maximizations.

Anyway, I enrolled in Economics and one of the things I read was a brand new book, Hick's *Value and Capital* [7]. You know, after reading though the mish-mash like Marshall and things like that, suddenly there was this clear, well-organized view, you knew exactly what was happening. Just the sort of thing to appeal to me. There was a whole, messy, confused literature on capital theory; all those great debates between Knight and von Hayek and all that. And now here was just the idea of dated commodities and suddenly scales fell from your eyes. A simple idea like dated commodities made whole issues transparent.

But as I read Hicks, I could see there were things left out. I turned to this again when I returned from the War, which was really pretty much of a hiatus in any work I was doing – I was gone and very busy for about three and a half years. I had done all my examinations before I had left. So now it was just a question of my thesis. I decided to take *Value and Capital* and redo it properly. I could see all kinds of specific points that were of concern. I wanted to combine it with Samuelson's stability theory, which he had developed in the meanwhile, the papers on dynamic stability in '41 and '42 [8]. Maybe I would add some stochastic elements to the story because as a student of probability and statistics theory I could see noise in the system. Well, it was a lifetime of work, really; it was a very unrealistic thesis.

Hotelling was primarily interested in statistics at this time and then he left for the University of North Carolina. And Abraham Wald wasn't interested in Economics anymore, either. The one I was closest to was Albert Hart, who was regarded then as a very promising theorist, but somehow wasn't able to do what he was capable of. Now people haven't even heard of Albert Gailord Hart. He had a good analysis of flexibility in a Festschrift for Henry Schultz [9], another figure who has faded, but I was never impressed by Henry Schultz. Hart considered a problem where you're thinking of buying a durable machine and you're uncertain as to the second period output; the trade-off is between a first machine that would be beautifully optimizing if you knew exactly the output but is not very good at slightly different outputs and a second machine that has costs that are fairly uniform along a wide range. The second machine might sometimes be preferred. He gave a sequential analysis; the idea that your choice today can be dependent on your uncertainty about tomorrow. Elementary as that point may seem, it just hadn't been expressed anywhere; it was very revelatory and came out in this Festschrift for Henry Schultz [10], published in '42, '43 or '44.

Hart was friendly and respectful but not very mathematical. One of the things he brought to my attention was that firms are, after all, multiowner objects. It is true that all the owners are interested in the same thing, maximizing profit; however, from a Hicksian point of view, the owners might have different expectations. Then their recommended investment policies today would be different. Each one, trying to maximize the same thing, expected profit, would nevertheless have a different choice. Now, from a modern point of view, we would probably take a different position. In fact the whole idea on my part was wrong; I didn't take account of the very simple point that owners could sell their stock. I didn't think about until I read Modigliani-Miller [11] years later and realized that my whole attack had been wrong. Those who were most optimistic about the firm would, in effect, buy it out from everybody else, who would do something else with their money. So, in fact, in that context, the voting paradox is irrelevant. But I did not think of it that way, I thought of owners as glued to the firm, and just did not think about the stock market. Within that context I thought, well, how would they decide between two actions. A reasonable thing is to assume it goes by majority vote, a majority of shares, of course. I started writing this down and it occurred to me that I have a preference of the firm defined by the statement that a majority prefers investment A to investment B. Then, from my background, a natural question is, is this relation transitive? Well, it didn't take more than a couple of tries to see that this is not true.

The minute I saw it, I thought: This must be well known. In fact, I thought I might have seen it before. I have no idea from that day to this what I could have seen. However, it is the sort of thing which would appear maybe even in a puzzle page of a newspaper. It could have appeared in some quite trifling way. Anyway, I thought I'd seen it before and I didn't think it was major; all I thought was it was a nuisance because it was spoiling my theory. I ended up trying to develop a theory on the basis of maximizing profits weighted by share numbers, using that as a maximand. Later I gave up the whole thing because it seemed very unwieldly and didn't cohere.

JK. Gave up that section or the whole dissertation?

KA. The whole dissertation. Large parts were in no way novel. At best, I learned some ideas about myopia in investment, things that in later years I pursued. At one point I had about thirty pages of outline of ideas and results but somehow it didn't send me. I felt somehow it was supposed to lead ultimately to empirical work. I was rather discouraged because I was spending quite a bit of time at this. It was a couple of years.

JK. Now comes the Hicks lecture . . .

KA. Intransitivity was something I had discovered, but it was not on my mind; it was something I had dismissed. Hicks gave a lecture during the winter of '46–'47. He had a very interesting idea. He was trying to find a definition of welfare inequality which was nevertheless consistent with ordinalism. What was meant by saying "individual A is better off than individual B?" Hicks' statement was the following: Suppose individual A prefers his own bundle to B's and individual B also prefers A's bundle to his own. Then Hicks would say A is better off than B. Of course, this definition surfaced again about twenty years later in the work of Duncan Foley [12], but I'd never heard of it before. I think I once found some indication that Trygve Haavelmo had had that idea and I've tried to find the reference again but I've never succeeded in tracking it down. Hicks presented this lecture and went through a lot of variations of this point but he said Joan Robinson had criticized it and he was a little worried about it. He said, of course, that people might be non-comparable. A could prefer his bundle to B's and B could prefer his own to A's – he recognized that. I thought about it sitting there and finally said, "Would you want this property of 'being better off than' to be transitive?" I defined what I meant by transitive. "Well, in the case, your definition won't satisfy that, because the comparison between A and B is based on A and B's orderings, the comparison between B and C is based on B and C's orderings, and the comparison between A and C is based on A and C's orderings. And it's possible to have A better off than B, B better off than C and C better off than A by this." So I guess somehow this idea of intransitivity fascinated me. Hicks said, "What? What? What?" I don't think he quite got the point. Hart was very quick; he was chairing the meeting, got the point and tried explaining it. Interestingly, Hicks never published this. Now all this may be beside the point, because you may not want transitivity even though it seems natural. In Foley's work, for example, this issue doesn't arise. Foley just said a point is fair if nobody is better off than anybody else. But this story does show intransitivity was bubbling around inside me even though I wasn't conciously aware of pursuing this line as a subject of research.

At this time, I received an invitation to the Cowles Commission. At first I postponed a move because I was trying to finish my Hicksian dissertation before I went there, but I finally settled on finishing it there.

JK. Wasn't it unusual then to leave graduate school before finishing your dissertation?

KA. You know, my knowledge of what was typical wasn't very good. The people I knew were at Columbia. I didn't know what was going on at Harvard or Chicago. It was really very provincial. In fact from what I now know the Columbia situation was unusually chaotic. One problem at Columbia, and I think it's true to this day, is that the sense of community among students and faculty is very weak; they're all dispersed. In particular the National Bureau of Research was a very strong organization and some of the leading Department members went off there, especially the great Wesley Clair Mitchell. So they weren't available. The Bureau was not near the University and so they were just simply physically somewhere else. One had the feeling, in fact, that they never talked to each other.

JK. What led to your invitation to Cowles?

KA. They came around and asked Wald and he recommended me. While he was primarily trained as a statistician, nevertheless he was interested in economics. There weren't many in that category and so they asked him.

Cowles was a funny kind of place because they were kind of a persecuted sect; the mathematical and quantitative emphasis was exceptional and distrusted. My salary was $ 3200 per calendar year – it was a calendar year appointment. Even by the standards of 1947, that wasn't very much.

JK. What were you hired to do?

KA. What they really wanted me to do was work on statistical problems but it was a free-wheeling place. At the moment the emphasis was on the development of the econometrics of large-scale models. So called "large": three equations, five equations. Larry Klein ended up with a 20 equation model. Now Tinbergen had even bigger models in his League of Nations study but this was simultaneous equations estimation which made the computational burden very much greater. I had some idea of using higher order approximations. Others had gotten the asymptotic distributions which were normal and I had been taking a course in Edgeworth-Cramér expansions which you get from higher-order approximations. These ideas were originated by Edgeworth and quite ignored; interestingly, Edgeworth authored quite a few new ideas in statistics, most of which were ignored and then rediscovered. Edgeworth had this method of Cramér rediscovered it. Actually, it was pretty high-powered mathematics and it really was probably beyond me. I knew how to do it, but mathematical estimates of the error term in the approximation were a very subtle and complex matter.

But I was there to do anything I pleased and I was very obviously interested in theory. There was a feeling that theoretical foundations were also an essential part. Finishing my thesis could fit into this.

JK. While you were at Cowles you worked on the single-peakedness result.

KA. I really spent a year there not doing much of anything, to tell you the truth. I wrote a few tiny papers, none of which amounted to anything. I was a great contributor to discussions: argumentative, finding exceptions, errors and counter-examples. But I really felt very discouraged. Once, at lunch, we were talking about politics, left parties and right parties, and I remember drawing on a piece of paper the idea that a voter might have preferences over the parties. It wasn't so much that I saw the ideas – it was the only way I could think about it. It was not that I thought why don't we represent voters as having preferences – as soon as I thought about the question, it couldn't occur to me there was any other way of doing it. So I wrote this thing down and started looking at the question of majorities. Its really hard to describe it. All I can say is, once you've seen it, its obvious; it takes an hour or two. If you ask the question, the answer is fairly obvious. I spent a day or two working it up as a formal proof. And in my usual way, I sort of stalled about a month on writing it up for publication. No, but that doesn't make any difference. I can't say a lost anything. It would just establish that I had the idea independently. But in a sense it didn't matter because within about a month I picked up the *JPE* and there's the paper [13] by Duncan Black that had exactly that idea.

The coincidence I regard as an extremely interesting point in the history of thought. It's an idea which could easily have occurred to Condorcet. It doesn't depend in any way on the development of mathematics in the last 150 years.

JK. Except to a sensitivity to the logic of relations?

KA. I suppose so. Well, let me put it this way. The logic of relations was worked out at great length in the latter part of the 19th Century. Let's say it depends on Boole and the idea of relations as ordered pairs – even that goes back to 1910 or 1911. But while the idea of relations as ordered pairs is a comforting idea, in the sense that you have a logical foundation, it isn't necessary. The fellow who developed a lot of the ideas about logical relations, rather sophisticatedly, was Charles Peirce, the philosopher and logician, founder of pragmatism around 1880–1910. It was picked up by a German named Schroder who in good Germanic fashion wrote three large volumes [14] around 1890. All the apparatus, all the sensitivity was there. If you needed more, the *Principia* surely supplied all that was needed. Nobody asked that question; that's all I can say. Black of course, had been building up to it; he really did have the idea of voting as a mechanism.

JK. Had you read anything of Black's before this paper?

KA. I honestly can't tell you. The stuff before was awfully formal and obvious. This was the only paper of his I seriously regard as having some excitement in it. So, anyway, there was my third encounter with orderings. But that really developed out of amusement.

Then that summer I went to Rand Corporation – again through sheer accident. My wife, who I met as a graduate student in Chicago, had previously worked in the Agriculture Department. She'd arrived there as a clerk and became a professional, a statistician. Her boss was a very distinguished mathematical statistician named M. A. Girshick. So I was friendly with Girshick who had gone to the Rand Corporation when it was started. The Air Force needed someone to tell them what was going on in the world, so they took all these wild characters and unleashed them. Girshick was one of those invited to go out there and he commenced to spend a couple of years. He often visited Chicago. He had been in contact with the Cowles Commission anyway, because some of his work in multivariate analysis really was very close mathematically, more than mathematically, close conceptually, to simultaneous equation estimation. He was giving advice to them; in fact, he had some ingenious ideas. He contributed a good deal to the development of the limited information method and was never really given full credit for that. Anyway, Girshick had this connection with Cowles independent of us, but when he came to Chicago he visited us.

One of the things Rand was doing was inviting large numbers of visitors for the Summer so Girshick urged me to come. Summer in Santa Monica didn't seem like a bad idea to me and it turned out to be far more intellectually exciting than anything I had planned because the halls were filled with people working on game theory. Everybody was fooling with zero-sum games, how to calculate them, the fundamental definition of the concepts; it was work at the conceptual level and at the technical level.

JK. Was game theory something you'd studied before you went to Rand?

KA. Not really. I mean I knew about the book and had been vaguely pecking away at it, but I hadn't really studied it at all carefully. It was not a big topic at the Cowles Commission, although Marschak had written a review of it.

Anyway, Olaf Helmer was among those who had been brought to Rand; a philosopher. There were several people there as a matter of fact who were basically philosophers; Abraham Kaplan was another. Helmer said to me one day: "There's one thing that disturbs me." They were taking game theory and applying it especially to Soviet-U.S. relations: diplomatic conflict, potential tactical situations, war. However, the payoff functions were defined in terms of utility functions, as Von Neumann and Morgenstern argued in their appendix, and these were derived on the basis of the individual. The trouble was, the Soviet Union and the U. S. were not individuals. What is the meaning of this? You've got to give him credit for proposing the problem.

Now I hadn't spent a lot of time or attention on welfare economics. I had really been trying to work on descriptive theory and general equilibrium theory considered as descriptive, rather than as normative theory. But I did read. One of the things I did when I was supposed to be working on my thesis was read and read and read. It was the typical thing I was doing. I still do it to this day. I'm supposed to be doing something and I find myself picking up something alledgedly relevant and reading it. You pick up a lot of information that way at times. I had read Oscar Lange's expository article [15] on the foundations of welfare economics. Lange was extremely clear. It was only afterwards I began to feel his clarity was purchased at the price of depth. He set forth very clearly the conditions for a Pareto Optimum. But then he referred to the fact that one could consider maximizing a welfare function. You started off with $U_1, \ldots, U_n$, utility functions of individuals, and you want to maximize $W(U_1, \ldots, U_n)$ but then there are a whole subset of the maximizing conditions that don't involve W – those are essentially the conditions that define Pareto Optimality. If I recall correctly, he was quite clear on distinguishing these concepts but he did have this W function and he gave a rather casual reference to Bergson. By this time, Samuelson's *Foundations* had been published and he gives a very full account of welfare economics in Chapter 8 which he bases on Bergson's paper.

So I gave a quick reply to Helmer: "Economists have thought about that and its really explained by Bergson's social welfare function." "Oh, is that so," he said, "Why don't you write it up? I think it would be nice for us all to have an exposition of how the Bergson social welfare function settles this."

JK. So you just started to write this up.

KA. Well, of course, I dropped the U's which I never liked because I knew the U's were just disguises for R's for preference relations. I thought, while I was at it, I'd do an exposition starting from just the orderings. Then I started musing about what information is conveyed. Welfare comparisons could be regarded as a series of pairwise votes and I was obviously interested in elections so that just seemed like the natural language to use. One natural method of taking a bunch of R's and putting them together would be by pairwise comparisons by majority voting. And I already knew *that* was going to lead to trouble! So I figured, well, majority voting was just one of a very large number of possibilities, you just have to be more ingenious. I started to write various possibilities down.

JK. For example . . .

KA. I'm pretty sure the Borda method came to my mind, because that was a very well known method. I didn't know it was Borda, you understand.

JK. It was something you'd encountered before?

KA. That was very well known; it was a widely used custom. A club might do that. It was something that was done in practice.

JK. Did you ask any political scientists at Rand about voting procedures?

KA. I don't think at first they did have any political scientists. I don't know anybody I would have regarded as a political scientist. I don't think Rand was interested in traditional political science. Later I think they had more traditional political scientists but even when they did they had people interested in area studies, Soviet experts. I don't think they ever had theoretical political scientists.

JK. So this experimentation was totally isolated. You didn't ask anybody about rules.

KA. Right, but that also reflected me. It seems to me I was trying at some point to systematically go through all possible rules. I took some examples and then consider related examples. Beyond that, I can't tell you which rules I explored. I did grasp that, at some point in this procedure, that part of the point was I was only using information on the alternatives under consideration. But then that struck me as a very natural thing to do.

JK. Very crucial.

KA. That turned out to be *very* crucial. In fact, if I had realized how crucial it was, I might have been more disturbed. It seemed very natural. Afterwards, when I formalized it, I saw the importance of it. Now it was obvious enough that if you let *one* person make the decision, there isn't any particular problem. So I was assuming non-dictatorship. But, I think it was a lucky thing, I assumed non-dictatorship in a very strong form. You know at the start I was only looking at triples, because that had the essential problem. I was assuming there wasn't any dictator or any pairwise choice. And then you see, you need two or more individuals to be decisive. A little calculation shows that you can always produce orderings that violate this. It really is the Condorcet paradox restated. And then when I tried to extend it to more than triples, I retained this postulate.

So, in this first version, which I did show about a week later to Abraham Kaplan, there was this idea of showing that these conditions were incompatible. But I thought I had really put a lot of emphasis on the assumption that there was no dictator on any pair; so, I thought, well there must *be* a solution if you allow different dictators on different pairs. This seemed to be absolutely crucial to the argument, and I thought about this for awhile. I thought it would be easy to produce an example where you have a different dictator on every pair. Now I didn't think that was a suitable solution anyways, I thought an assumption of non-dictatorship was a correct assumption, even on one pair, so I wasn't too disturbed. But I thought, to clean out the exposition, to show exactly what was meant, I ought to produce an example where you have different dictators on different pairs. But no, if there's transitivity that means dictatorship sort of propagates itself. And finally one night when I wasn't sleeping too well, I could see the whole proof, you know after playing around with it for awhile. And that was a couple of weeks later, that I had the idea that the non-dictatorship condition could be stated in this much weaker way, that the whole ordering can't be determined by one dictator.

JK. What did you feel at this point?

KA. I felt this was very exciting. I thought "*This* is a dissertation." You know, its a funny thing. One of my problems had been feeling that one has to be serious and every time I'd thought about these voting questions they seemed like amusing diversions from the real gritty problem of developing a good descriptive theory. And in some sense I still have a little bit of that feeling. But when I got the result, I felt it was significant. I really did. It clearly didn't conform to my preordained ideas about what was significant. I would have said a priori if somebody told me about this, my temptation would be to say, "Well, that's very nice, but what importance is it?" But when I did it, I felt, yes, this is something. This was at least asking some very fundamental questions about the whole nature of social intercourse and particularly about legitimation of collective action.

This wasn't just a technical issue in game theory. The technical and the philosophical were intimately merged. My whole work in general, not only in this field but in others, has tended to deny the idea we can take off the technique and put it here and put the deep issues there. Some of the so-called technical issues are really of the essence of the so-called deep issues and you really can't separate them at all. Each one illuminates the other. In fact they fuse together and in some cases they're identical. And nothing can better exemplify this than social choice theory where the central issues and the technical issues were identical.

JK. Let's go down the list of some of the names you've acknowledged and tell us what they contributed. First the people at Rand: Abraham Kaplan.

KA. Well, he was one of the first persons to whom I showed the results. He was the only one who combined the philosophical side and at least some of the technical capacity to appreciate this. I don't recall getting anything specific from him, just an appreciation that it really was important.

JK. Youngs.

KA. I don't know why I thanked Youngs. Youngs was a mathematician with whom I discussed some issues of preference orderings and the like. I had been in close contact with him and I discussed some other aspects about preference orderings, really about individual preference orderings. I felt a kind of general intellectual debt.

JK. David Blackwell.

KA. Again, well, Blackwell was a genius. He and I worked very closely on other matters. He, Girshick and I wrote a paper on sequential analysis [16]. He did contribute one other thing. There was this chapter [in *Social Choice and Individual Values*] which has never been followed up. I was raising the question about if the orderings were restricted in some ways, when does the paradox exist. If you go to the extreme of single-peakedness we know the answer. So the question is, supposing you have *some* restrictions on orderings but there is a lot of freedom left. I had a result, where the technical point was when could you extend a quasi-ordering to a full ordering and it was Blackwell who told me about Szpilrajn's Theorem. However, unfortunately, my proof is not correct, because it suffers from the problem that Blau pointed out. I suspect the theorem is correct or some theorem like it is correct but nobody's ever stated it and I've never gone back to it.

JK. J. C. C. McKinsey.

KA. McKinsey was a very interesting fellow. He was the one who educated all of us to what game theory was all about. So the influence was indirect, but in a way it was there. He was a beautiful expositor. He was a logician of considerable power and had done some work earlier on the formalization of logic; he was a disciple of Tarski's. The whole game theory ambience, and therefore in particular McKinsey – and Blackwell for that matter on the technical side – were influential in setting the whole tone to this.

JK. The next names are from Chicago: Tjalling Koopmans, Herbert Simon, Franco Modigliani, T. W. Anderson, Milton Friedman, David Easton.

KA. The exposition of the book was developed in the next year back in Chicago. I presented the material over a number of seminars. I was grateful to these people because they thought it was a good idea, encouraged me and asked good questions; parts of the book are making clear points they found obscure.

Easton was a little different. He was the first political scientist I talked to about this. He gave me the references to the idealist position which was sort of the opposite idea. In a way the idealist position was the only coherent defense that I could see in political philosophy. It wasn't a very acceptable position, but it was the only one that had at least a coherent view of why there ought to be a social ordering.

JK. Why did you call it a "Possibility" Theorem?

KA. That was Tjalling's idea. Originally I called it an impossibility theorem, but he thought that was too pessimistic! He was my boss and a very sweet man, so I changed it for him.

JK. There was a meeting of the Econometric Society where you presented these results.

KA. I guess I must have presented it at the December, 1948 meeting.

JK. Who was there and what was the reaction?

KA. I remember Larry Klein was in the chair and Melvin Reder was reading another paper at the same session. My recollection is that there were 30 or 40 people in the room. I distinctly remember that in the audience was this contentious Canadian, David McCord Wright, who objected because among the objectives, I hadn't mentioned freedom as one of the essential values in social choice and apparently he went out of the room saying that Klein and Arrow were communists – this was quoted to me at least by Kenneth May who was also present.

I thought under the circumstances, I got a pretty good reception. I don't think anybody said "We've seen a revolution before our eyes," but it *was* taken as a serious contribution. I wonder why it was accepted so well. There really was no resistance. It made my reputation.

There had been, of course, a fair amount of controversy about the foundations of welfare economics, beginning with papers by Harrod, Hicks, then Kaldor, then the long chapter in Samuelson's *Foundations*, then Scitovsky in 1941 with intersecting community indifference curves [17]. So unease about the foundations of economic policy were there. So the debate was serious – people were already concerned about these things.

Right after the summer I developed this, on the way back to Chicago, I stopped at Stanford to be interviewed for a job. Girshick had meanwhile moved to Stanford to contribute to starting a Statistics Department there. He was their star and he wanted me to join him. The Economics Department there had already in fact made me an offer a year earlier.

JK. Based on?

KA. What happened was due to Allen Wallis, who was recently Chancellor of the University of Rochester and is now, at the age of 75, Undersecretary of State for Economic Affairs. Wallis had been a Professor at Stanford before the war, he was the first really major appointment they ever made. He didn't return after the war, but he was highly regarded and apparently they asked him for recommendations. He had worked with Wald during the war and Wald had spoken about me based on my work as a student at Columbia. It was very common at Stanford to appoint Assistant Professors who didn't have a PhD; they assumed we would finish. I was appointed without a PhD. In fact, I got tenure without a PhD.

JK. Really?

KA. Well, I'm being a little technical, but my statement is technically correct. In those days you couldn't get your degree until your dissertation was printed. So I had these theorems and then sent my changed proposal into Albert Hart, who got kind of excited about it. I defended the thesis in January of 1949. Stigler was on the examining committee, Bergson had come to Columbia in the meantime and they put him on the commitee. Of course, Bergson was asking some searching questions but was very fair and did have a high opinion of me. But I had no real interaction – I sent in my dissertation and got it approved.

Hart was immediately enthusiastic. He said, "I don't really understand it fully, but it sounds like you're dealing with very important issues," and I've heard later that around Columbia it was held to be an exciting event. Of course I had been regarded a kind of a star student. In fact, one of the things that had worried me was whether I was just an eternal student.

Stanford had a custom where all initial appointments were for one year. They kept this idea since they frequently hired people without even interviewing them – because of the geography. Even Moses Abramovitz who transformed the Department was hired as a full Professor for just one year just before I came. The jet has ended all this. So I came as an Assistant Professor on a one year appointment. My thesis had been approved, But I couldn't get a degree until it was printed. This was just about the end of that era. I must have been one of the last people to come under that rule. In fact, while the thing was in the process of being printed, I received a notice that if I submitted a typed manuscript, I could get my degree immediately. But that was kind of expensive, to have somebody retype that all up. I really lost about a year on my degree by deciding to go ahead on the printing. But then they gave me tenure on the basis of this unpublished dissertation. You couldn't do it today. It would never be approved today.

Then it is interesting – the reception question. Hart, who didn't work in this line, was very enthusiastic; he had spoken, I gather very well, around the Columbia faculty. The people at Stanford were very impressed; essentially all I had for them to see was this work – I hadn't done anything else except trivial stuff. They were so impressed that by the end of the interview day – within a day they were ready to make me an offer. So it's interesting to get this reception from all sorts of people not logically trained, not mathematically trained. And when the book came out, it made a great success. It is a little puzzling – at the time I took it as that's what happened. But in retrospect I sort of wonder why.

JK. Let's go to Blau's discovery of a mistake in your proof [18]. *That was eight years later; was that a great surprise?*

KA. Yes, it certainly was. Blau was working that year at Stanford and showed it to me. I was surprised, but I knew right away that a universality assumption would correct things. It had seemed obvious to me that the non-dictatorship property was hereditary, but it wasn't. I still think there is a better correction than that one, but I've never really gone back to work on it.

Blau's was a very, very nice result. It didn't obviously change the basic impact, but it did show my little attempts at generalization didn't work. It's interesting to see how easy it is to make a mistake on things that seem so airtight.

JK. Let's leave the origins now. Over the succeeding 40 years, what were the most important developments in social choice theory?

KA. Some of the work that has cohered around the original question is mathematically interesting but not very relevant to the original field. The literature that depends on small numbers of alternatives is in this category. I think the alternative space must be taken to be very large.

I also have qualms about results like those of Kirman and Sondermann [19]. What do we learn from Kirman-Sondermann exactly? When you have an infinity of voters, then the axioms as I wrote them become consistent and you can produce voting systems. But they are consistent because in some sense the dictator has a different meaning; banning a dictator is no longer enough. As it turns out, if you have a sequence of decisive sets, each of which is properly contained in its predecessor, the intersection of the whole sequence is empty, but everyone in that sequence is then less decisive. In some sense, the spirit of non-dictatorship ought to rule that out.

Incidentally, there's a recent result I haven't had a chance to study by a former student of mine, Alain Lewis, who says if you confine yourself to recursive functions then the voting paradox occurs even with an infinite number of voters; in the strict sense, even with just the ordinary axioms. The examples Kirman and Sondermann use are non-constructive; something with cofinite sets is not something you can actually construct – you just show it exists. But since they are only examples, that's not a proof that there isn't a constructive procedure. But Lewis says he's given a proof and I have to study it. If I can understand it.

JK. What about the Gibbard-Satterthwaite [20] *results?*

KA. Gibbard's work was a bombshell. That was very exciting. I didn't know about Satterthwaite's work for a couple of years, but it was very much the same thing. I had taken the liberty of abstracting from manipulability in my thesis and I never went back to that issue. What's surprising is not really that there is an impossibility of non-manipulability, but that the issues should be essentially the same. That strikes one as a remarkable coincidence.

I still find it surprising and feel that we might not have the right proof. Somehow you feel that if you had the right proof it would be obvious. But then I thought that about my work, too. My impossibility theorem ought to be totally obvious when looked at the right way. Yet every proof involves a trick. Maybe not a big trick; I don't think it's a mathematically hard theorem. But somehow if you had the right way of approaching it, it should be trivial. Yet, no matter how you present the proof, and they're all pretty close to equivalent, its not yet trivial. For example, when ultrafilters came in, I thought, Aha!, this is a beautiful way of showing it. But it turns out that to prove the decisive sets form an ultrafilter involves essentially all the original calculations.

JK. Still, it's a nice approach conceptually.

KA. I don't know. I'm less convinced than when I first saw it. It has the advantage of referring to a known body of knowledge. But this is a body of knowledge which is somewhat technical. You're bringing in a fair amount of technical apparatus and it ought to pay for itself somewhere, if I may use an economic approach. It ought to pay for itself in making the proof trivial. But in fact you need just about every step in the original proof to show that the issue *is* one of a fixed ultrafilter. So therefore, why bring in all this apparatus? I was a little surprised by how little you get from all that apparatus. I can't help feeling there's some way out of it. In the same way, I always feel the Gibbard-Satterthwaite result should be more transparent than it is. But maybe it can't be done.

JK. What about Sen's Paretian liberal approach [21]; *does that interest you?*

KA. I thought that was stunning and penetrating to a very important issue. But . . .why do we have rights? What I am after all is a kind of utilitarian manqué. That is to say, I'd like to be utilitarian but the only problem is I have nowhere those utilities come from. The problem I have with utilitarianism is not that it is excessively rational, but that the epistemological foundations are weak. My problem is: What are those objects we are adding up? I have no objection to adding them up if there's something to add. But the one thing I retain from utilitarianism is that, basically, judgements are based on consequences. Certainly that's the sort of thing we do in the theory of the single individual under uncertainty; you make sure utility is defined only over the consequences. I view rights as arrangements which may help you in achieving a higher utility level. For example, if you are much better informed about a certain choice, because it's personal to you and not to me, I don't really know anything about it, I should delegate the choice to you.

JK. You don't want to allow preferences over processes?

KA. Well, one of the things I fear is emptiness. You put preferences over enough things, then anything that happens can be defended. It destroys the idea of discourse. Of course, it is a delicate issue, you can always say, of any particular process that it is specially privileged. You could take Nozick's point of view; you can have an absolute preference about certain processes. For example, we have a property system; if you and I make an agreement about anything within our property rights, that just fixes it, period. Now I've got to admit Nozick's courage is good. Suppose somebody invented a cure for cancer and allowed it to be used only at an extremely high price. Nozick says: No problem. Most everybody else would regard that as a fatal counterexample, but Nozick has the courage of his convictions. But that's a strong example of preference over processes. Most of the people who are advocating rights are very different, like Dworkin. They tend to support so to speak left-wing rights rather than right-wing rights, but once you grant that, who settles what rights are legitimate? The consequentialist view – I won't say that fully settles it either, but at least you have something to argue about. So this is why I'm a little unsympathetic to the rights issue – everybody just multiplies the rights all over the place and you get total paralysis.

Consider the consenting adult example – say homosexuality – and think about the concept of externality. Now why do we say intercourse among consenting adults should be allowed? One argues because there's no externality. But if I care, there *is* an externality. I actually allude to this even in the first edition of my book. It's just a rhetorical passage and doesn't enter the logic, but I mention that the concept of preference is just what everybody thinks their preferences are. Different people might have different ideas of externalities. I took the view that all preferences count. From the logical point of view, it doesn't matter; if you purify the preferences by rejecting the nosy preferences, the theorem applies to whatever is left. There is, of course, a technical problem in systematically combing out inadmissible preferences. Transitivity says you can't just look at separate preference pairs, you have to look at the whole system. That's what Gibbard's paper is really devoted to. Gibbard is not totally convincing, because there are some arbitrary choices in his elimination procedure; he doesn't make it compelling that his is the only way of doing it. It's just *a* way. It looked pretty devastating but it eliminated more than was really necessary.

I'm quite puzzled. People really care about consequences as they see them. If I'm really offended because people are seeing obscene material, well, I'm hurt. I really am hurt. I'm hurt just as much as if somebody blew smoke in my eyes – or whatever your favorite form of pollution is. Indeed a lot of people probably care much more. I really find it difficult to decide.

Unless somebody produces a logic of rights in terms of which we can *argue*, I really find the whole issue is unfocused. The reason why it is compelling is that there are at least some cases where we do feel strongly about the rights. It's not clear you can always reduce those to utilitarian considerations like information.

JK. One last question. What outstanding problem in social choice theory would you most like to see solved?

KA. Well, if I had to pick just one, it would be reformulating a weakened form of the independence of irrelevant alternatives which stops short of just dropping it completely. There are a lot of arguments used today, extended sympathy, for example, or the relevance of risk-bearing to social choice as in Harsanyi or Vickery [22], that do involve, if you look at them closely, use of irrelevant alternatives. Suppose I'm making a choice in Harsanyi's story among totally certain alternatives. I somehow use preferences among risky alternatives as part of the process of social decision. We use a chain of reasoning that goes through irrelevant alternatives. It seems quite open to acceptance, not at all unreasonable, that these are useful. I would not want to rule out in an argument, a line of reasoning which goes through a chain of transitivity via an irrelevant alternative. And yet I don't want to be in the position of saying, well the whole thing depends on the whole preference ordering. My current feeling is that that is the most central issue – the most likely way of really understanding issues.

JK. Are you anticipating that if you allow chains of transitivity over irrelevant alternatives you will obtain a "good" social choice procedure or are you expecting a deeper impossibility theorem?

KA. I'm expecting – no, let me put it more cautiously – I'm *hoping* for a possibility result.

References

1. Tarski A (1941) Introduction to logic and the methodology of the deductive sciences. Oxford University Press, New York. Appeared originally in Polish, 1936, and was translated into German, 1937
2. Kenney JF (1939) Mathematics of statistics. Van Nostrand, Wokingham (2 vols)
3. Hotelling H (1932) Edgeworth's taxation paradox and the nature of demand and supply functions. J Polit Econ 40:571–616
4. Hotelling H (1935) Demand functions with limited budgets. Econometrica 3:66–78
5. Popper K (1938) A set of independent axioms for probability. Mind 47:275–277
6. Wold HOA (1943, 1944) A synthesis of pure demand analysis, I–III. Skandinavisk Aktuarietidskrift 26:85–118, 220–263: 27:69–120
7. Hicks J (1939) Value and capital. Clarendon Press, Oxford
8. Samuelson PA (1941, 1942) The stability of equilibrium. Econometrica 9:97–120; 10:1–25
9. Lange O, McIntyre F, Yntema TO (eds) (1942) Mathematical economics and econometrics: In memory of Henry Schultz. University of Chicago Press
10. Hart AG (1942) Risk, Uncertainty and the unprofitability of compounding probabilities. In: Lange O, McIntyre F, Yntema TO (eds) Mathematical Economics and econometrics: In memory of Henry Schultz. University of Chicago Press, pp 110–118
11. Modigliani F, Miller MH (1958) The cost of capital, corporation finance, and the theory of investment. Am Econ Rev 48:261–297
12. Foley D (1967) Resource allocation and the public sector. Yale Econ Essays 7:45–98
13. Black D (1948) On the rationale of group decision making. J Polit Econ 56:23–34
14. Schroder E (1890–1905) Vorlesungen über die Algebra der Logik (3 vols). B. G. Teubner, Leipzig

15. Lange O (1942) The foundations of welfare economics. Econometrica 10: 215–228
16. Arrow KJ, Girshick MA, Blackwell D (1949) Bayes and minimax solutions of sequential decision problems. Econometrica 17: 213–244
17. Scitovsky T (1942) A reconsideration of the theory of tarifs. Rev Econ Stud 9: 89–110
18. Blau J (1957) The existence of social welfare functions. Econometrica 25: 302–313
19. Kirman A, Sondermann D (1972) Arrow's theorem, many agents and invisible dictators. J Econ Theory 5: 267–277
20. Gibbard A (1973) Manipulation of voting schemes: A general result. Econometrica 41: 587–601
Satterthwaite MA (1975) Strategy-proofness and Arrow's conditions. J Econ Theory 10: 187–217
21. Sen A (1970) The impossibility of a Paretian liberal. J Polit Econ 78: 152–157
22. Harsanyi JC (1955) Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. J Polit Econ 63: 309–321
Vickrey WS (1945) Measuring marginal utility by reactions to risk. Econometrica 13: 319–333