Wulf Gaertner

Kenneth Arrow's impossibility theorem stretching to other fields

Article (Accepted version) (Refereed)

DOI: 10.1007/s11127-018-0503-y

Kenneth Arrow‘s impossibility theorem:

Stretching to other fields

Wulf Gaertner

University of Osnabrück

and

CPNSS, London School of Economics

Final version accepted January 22, 2018

Abstract

Arrow’s impossibility result not only had a profound influence on welfare economics but was, as this paper shows, also widely discussed in philosophy of science and in the engineering design literature.

1. Introduction

Kenneth Arrow's impossibility result had an enormous impact on welfare economics and finally led to a new area called "social choice theory" and the launching of a new scientific journal with the title *Social Choice and Welfare*. Arrow's seminal work on the (non-)existence of a social welfare function also had a profound influence on philosophical writings about justice and equity. The list of these contributions is long. Only a few should be mentioned here. First and foremost in my enumeration is John Rawls's *A Theory of Justice* (1971), Thomas Scanlon's *What We Owe to Each Other* (1998) and Amartya Sen's *The Idea of Justice* (2009), but reference also should be given to contributions by Stephen Strasnick (1976), Patrick Suppes (1966), Lars-Gunnar Svenson (1977), and Larry Temkin (1993), among many others. However, as the title of this paper indicates, I shall not review this strand of literature any further, all the more so because that path has been well explored during the last three or four decades.

What I wish to do in what follows is, first, to take a closer look at another area within philosophy, namely a discussion that had its beginning in Thomas Kuhn's famous statement that there is no neutral algorithm for theory choice and, second - at first sight rather surprising, for me at least - a discussion of the relevance of Arrow's theorem in engineering, where the issue is how to make rational choices among alternative design options. This short summary already indicates the structure of sections 2 and 3 of this paper. A few concluding remarks are gathered in section 4.

2. Kuhn's no neutral algorithm thesis

In order to evaluate and compare alternative theories adequately, particularly in the natural sciences, Thomas Kuhn (1962, 1974, 1977) argued that theory choice should be based on at least five epistemic criteria or values, namely accuracy, consistency, scope, simplicity and fruitfulness. "Together with others of much the same sort, they provide *the* shared basis for theory choice" (Kuhn 1977, p. 357). However, different scientists can create more than one ranking of rival theories even if they agree that the evaluation should be done solely in relation to Kuhn's (1977, p. 358) criteria: "When scientists must choose between competing theories, two men fully committed to the same list of criteria for choice may nevertheless reach different conclusions. Perhaps they interpret simplicity differently or have different convictions about the range of fields within which the consistency criterion must be met. Or perhaps they agree about these matters but differ about the relative weights to be accorded to these or to other criteria when several are deployed together."

Kuhn argues that philosophers of science often have neglected the subjective elements that enter into theory choice. Algorithmic decision procedures that attempt to solve the theory choice problem often presuppose "that individual criteria of choice can be unambiguously stated and also that, if more than one proves relevant, an appropriate weight function is at hand for their joint application" (Kuhn 1977, p. 359). The subjective component, "idiosyncratic factors dependent on individual biography and personality", finally led Kuhn (1977, p. 361) to the conviction that there is "no neutral algorithm for theory choice", or, because of the role that subjective factors play in the evaluation procedure, that there is no unique algorithm, but several algorithms instead.

It was precisely this view (or interpretation of Kuhn's position) that led Samir Okasha (2011) to resort to instruments from social choice theory and to Arrow's impossibility theorem in particular. Okasha interprets Kuhn's five criteria from above as "individuals" who have their own preference ranking over alternative theories. So if R_i is the preference order of individual i over a finite set of alternatives X or, in the current context, the rank-order of criterion i over a set of given theories, then the issue is to aggregate any logically possible profile of orderings R_1, R_2, …, R_n into a unique ordering over X, which is exactly how Arrow set up the social choice problem.

One should pause for a moment in order to become aware of the fact that on the path from Kuhn to Arrow, Okasha introduced a non-trivial twist. While Kuhn's position was that no unique algorithm exists because of various subjective components within the evaluation procedure so that several or even many algorithms may be available for choosing among alternative theories, Arrow's impossibility result that Okasha invokes says that *no algorithm at all* exists that meets certain standards of how an aggregation of preferences should proceed. Okasha (2011, p.93) is, of course, well aware of this "fact" when he says that "where Kuhn saw an embarrassment of riches, Arrow tells us that there is nothing at all."

Once this twist in the interpretation of Kuhn's position is accepted, the following statement by Okasha (2011, p. 94) makes sense: "Since the Arrow-style impossibility result threatens the rationality of theory choice, and thus of science, it would be nice if there were a way out." And very clearly, as welfare economists know, various possibilities have been suggested in the social choice literature and have fairly recently been the object of discussion, controversially of course, in the debate on theory choice by philosophers. Most of these proposals are closely related to Arrow's conditions that the aggregation procedure should fulfil, namely full rationality of the overall preference relation, unrestricted domain, weak Paretian orderings, independence of irrelevant alternatives and the non-dictatorship requirement. Okasha discusses in particular a justification of the non-dictatorship condition since one may argue that among the criteria that Kuhn formulated, one of these might perhaps have absolute priority. The author considers a possible dictatorship of accuracy, but finally dismisses that idea. He also asks whether, perhaps, unrestricted domain can be given up since tradeoffs or correlations may exist among several of Kuhn's criteria. It may, for example, be the case that greater simplicity is reached by a sacrifice of accuracy so that an inverse relationship is possible between both values. Again, that path is dismissed by the author as a general way out of Arrow's negative result.

It is interesting to note that Okasha finally follows Sen's (1977) proposal to enrich the informational basis of the aggregation procedure. Just to remind the reader, Sen suggested that the underlying utility information be broadened or enhanced. Then interpersonal comparisons of, for example, utility levels are made possible so that Rawls's (1971) difference principle, which looks at individual positions under alternative social states, can reasonably be applied. One example that Okasha discusses at greater length in this context is the Bayesian approach to scientific inference.

Consider several rival hypotheses T_i, $i \in \{1,\ldots, m\}$, and a body of evidence E. The Bayesian approach assigns two scores to each theory T_i. The first is the prior probability $P(T_i)$, and the second is the conditional likelihood $P(E|T_i)$. Various ways of combining these two criteria into a decision rule are possible. Bayesians argue that one should multiply the prior by the conditional likelihood, namely to consider the quantity $P(T_i) \times P(E|T_i)$. Both $P(T_i)$ and $P(E|T_i)$ imply a ranking of theories and thus constitute a "preference" profile. The problem of deriving an overall ranking of theories as a function of these two orderings is an example of the aggregation problem. Okasha speaks in this context of a Bayesian theory choice "functional" and refers to Sen's (1970, 1977) concept of a social welfare functional. The Bayesian choice functional violates two of Arrow's conditions, namely unrestricted domain and the requirement of ordinal non-comparability. The Bayesian function is defined on the unit interval [0, 1] with the restriction that the sum of prior probabilities with respect to rival hypotheses (theories) be smaller than or equal to 1. What is more important, this function assigns numbers, namely probabilities that are absolutely measurable. Okasha argues that the fact that more than ordinal non-comparable information is being used is the primary reason why the Bayesian function is able to satisfy the other Arrovian requirements and thus generate a possibility result and, therefore, an escape route.

Jacob Stegenga (2015, pp. 273-274) finds this escape route to a Bayesian theory choice algorithm to be mistaken: "…[T]he Bayesian criteria are (merely) *post hoc* measures of the goodness of theories, whereas most of the theoretical virtues commonly discussed (namely Kuhn's five properties, W.G.) are substantive properties that *constitute* the goodness of theories." Stegenga

(2015, p. 269) thinks that Sen's suggestion cannot be fruitfully imported into theory choice since "in many cases, arguably most cases of interest in science, we only have ordinal and non-comparable measures of the support that a theoretical virtue provides to a theory." This position undeniably is in sharp contrast to a recent proposal by Wulf Gaertner and Nicolas Wüthrich (2016) in which the informational basis is enriched in such a way that inter-criteria comparability is made possible within a cardinal setup. It is a grading method based on a common language of qualitative verdicts.

As already indicated, the discussion in philosophy of science after Okasha's article is very reminiscent of the debate in social choice theory after Arrow had published his path-breaking monograph on collective choice in 1951. As we will see shortly, it is very similar to a discussion of optimal design decisions in the engineering profession.

Coming back very briefly to theory choice, Michael Morreau (2014, 2015) thinks that the assumption of unrestricted domain is inappropriate in relation to theory choice. He argues that the variety in the criterial orderings does not appear to be rich enough in order to warrant this requirement; he gives several examples from "toy science" and real science. Marcel Weber (2011), in contrast to Okasha, believes that the non-dictatorship condition should be abandoned in the case of theory choice. Fruitfulness in particular should be considered as a dictatorial criterion among the epistemic values. Finally, Davide Rizza (2014, p. 1852)) thinks that Arrow's impossibility result has no relevance for theory choice as soon as one exploits the "sequencing or betweenness information" that is contained in an ordinal profile. Here, Rizza refers to Don Saari (1995) who, in various publications, has been arguing for positional methods, the Borda rank-order scheme in particular. Arguments in favor of this aggregation procedure which takes "ordinal distances" into account will reappear in the next section.

3. From collective choice to engineering design

There are two situations in engineering design for which the Arrow theorem may or, perhaps may not, apply. The first one sees a single engineer who has to evaluate alternative designs (a new body for a particular type of car, for example) according to a finite number of criteria. At the end of the day, this engineer is expected to come up with a rational decision that systematically orders the alternative conceivable designs. This is a multi-criteria decision problem. The second situation depicts a group of several engineers, "having different responsibilities for different features of the design – e.g., structural integrity energy efficiency, control system robustness, safety, effectiveness of user interface" (Franssen and Bucciarelli 2004) who will value design alternatives differently, based on their individual preferences. This is a team decision problem. Both situations call for an aggregation procedure.

Maarten Franssen (2005, p. 42) believes that Arrow's theorem is very relevant and immediately applies to both kinds of decision problems: "The relevance to engineering design of the famous impossibility theorem for social choice … is receiving increased attention in the literature. Judgments on the importance of the theorem differ, however, depending partly on how the theorem is seen to apply to engineering design." At the end of his article, Franssen (2005, p. 55) asserts that "engineering design methodology could profit from the recognition of the structural identity of multi-criteria decision problems and social-choice problems in that it would provoke a closer look at the way 'solutions' to these problems are conceived in the area of social choice and would lead to research into the merits of these solutions for engineering design." Franssen mentions in this context the majority rule and Kemeny's method.

Franssen discusses one aspect on which social decisions and multi-criteria decision-making may differ. In collective decisions, all voters are treated equally (the anonymity condition), while in multi-criteria decision-making "the case where all criteria are considered to be of equal importance would rather be the exception" (Franssen 2005, p. 45). It seems, however, that such unequal treatment can be incorporated easily into the Arrovian setup by counting a particular criterion more often than others or, perhaps even better, by introducing an additional criterion that is closely related in importance or functioning to the first one. The latter suggestion would avoid the issue of weighting potential criteria differently.

Michael Scott and Erik Antonsson (1999) argue that Arrow's negative result does not apply to the multi-criteria engineering design problem. "Engineering variables are almost always ordered on an external scale, and preferences for engineering requirements are commonly single-peaked around an ideal target. Indeed, nearly all engineering requirements are of one of three forms: less is better, more is better, or closer to a particular target is better" (Scott and Antonsson 1999, p. 224). This may be true in a variety of cases, but what happens when some criteria follow the dictum of "smaller better than larger", while others follow the maxim that "larger is to be preferred to smaller"? Table 1 depicts such a case (Gaertner 2016). Four alternatives characterized by their (increasing) power are available. The criteria are weight, power to weight and cost. While criteria 1 and 3 follow the former "philosophy", criterion 2 follows the latter. According to criterion 1, we obtain that *a* is better than *b*, which is better than *c*, which again is better than *d*. Criterion 2 prefers *c* to *b*, *b* to *d*, and *d* to *a*; and criterion 3 finds *d* best, then *a*, which is followed by *b*, with *c* being last. These rankings, taken separately, are very intuitive. Taken together, they cannot be arranged in a single-peaked fashion. Actually, two of the triple-combinations display a latin-square structure, namely (*a*, *b*, *d*) and (*a*, *c*, *d*). This shows that Scott and Antonsson's argumentation may not lead too far as a general way out of the Arrovian dilemma.

Table 1 about here

The authors assert that a practical difference exists between social choice and engineering design. Designs may have to fulfil constraints. "A maximum stress indicates the point at which a design breaks and fails; government regulations must be fulfilled or a design is not allowed on the market" (Scott and Antonsson 1999, p. 224). But do these constraints, we have to ask, necessarily lead to a structured set of preferences, in the sense of single-peakedness or any of the other domain conditions proposed in social choice theory (Gaertner 2001, 2009)?

Contrary to Weber's (2011) argument in philosophy of science, Scott and Antonsson do not think that non-dictatorship should be abandoned as a requirement in engineering design. "Dictatorship by one evaluation *criterion* is not a rational solution.… Some engineering cultures may appear to have a dictator in the form of a single decision-maker, perhaps a manager with ultimate responsibility for all decisions; however, decisions will still be made by considering several criteria” (Scott and Antonsson 1999, p. 221).

While Scott and Antonsson refer, as we have seen, to the fulfilment of a particular structure that criteria may follow when evaluating alternative design options, Clive Dym et al. (2002) provide arguments for a particular aggregation mechanism that they consider as appropriate in the judgment of designs. The authors call this scheme “pairwise comparison chart”. It is equivalent, as they show, to the Borda count, which violates Arrow’s independence condition. The reason for this equivalence is that the vector of ranks according to the Borda rule is the “aggregated version of pairwise voting” (Saari 1995, p. 156). To substantiate their proposal, Scott and Antosson (2002, p. 238) argue that “from a practical standpoint, both designers and teachers of design have found that pairwise comparisons appear to work well by focusing their attention, and by bringing order to large numbers of seemingly disparate objectives, attributes, or data points. In addition, these rankings often produce good designs.” So again, we find a similarity in reasoning between theory choice and decisions in engineering design, two fields that do not seem to know one another, as one does not find any quotations from the other literature in either area.

4. Concluding remarks

In this paper we have tried to show that Kenneth Arrow's impossibility result, which shook up major parts of welfare economics, also had a larger impact on fields of science that are quite far away from economics proper. At first sight, one would not have expected that Arrow's negative result would ever be discussed seriously in the philosophy of science and in engineering. But a second thought makes clear that the "Arrow problem" also manifests itself in those areas. It is interesting to see that both in the philosophy of science and in engineering attempts were made to demonstrate that Arrow's findings do not seem to have any deeper relevance for other fields – as there was in welfare economics after the appearance of his *Social Choice and Individual Values*. Paul Samuelson (1967, p. 42), for example, argued that Arrow's result is much more a contribution to the discipline of mathematical politics than to the theory of welfare economics.

It is also interesting to witness stunning parallels in the argumentation that emerged outside economics. In both fields that we were looking at, one can find arguments to give up the unrestricted domain condition, in both areas one also encounters arguments in favor of violating the independence condition and for considering a Borda-type aggregation procedure. All of that debate shows how deep and profound Arrow's seminal work on social choice is. To refer to Samuelson (1967, pp. 41-42) again, "…I must admit that my vanity as an economist is gratified that one of the soldiers in our regiment should have made a contribution of universal interest".

Acknowledgements. I am very grateful to an anonymous referee and the editor of this journal for several perceptive comments and suggestions.

References

1. Arrow, K. J. ([1951] 1963). *Social choice and individual values*, 2nd edition. New York: John Wiley.
2. Dym, C. L., Wood, W. H., & Scott, M. J. (2002). Rank ordering engineering designs: Pairwise comparison charts and Borda counts. *Research in Engineering Design*, *13*, 236-242.
3. Franssen, M. (2005). Arrow's theorem, multi-criteria decision problems and multi-attribute preferences in engineering design. *Research in Engineering Design*, *16*, 42-56.
4. Franssen, M., & Bucciarelli, L.L. (2004). On rationality in engineering design". *Journal of Mechanical Design, 126*, 945-949.
5. Gaertner, W. (2001). *Domain conditions in social choice theory*. Cambridge, UK: Cambridge University Press.
6. Gaertner, W. (2009). *A primer in social choice theory*, rev. edition. Oxford: Oxford University Press.

7. Gaertner, W. (2016). Aggregating qualitative verdicts: From social choice to engineering design. *Open Journal of Applied Sciences, 5,* 319-326.
8. Gaertner, W. and Wüthrich, N. (2016). Evaluating competing theories via a common language of qualitative verdicts. *Synthese, 193,* 3293-3309.
9. Kuhn Th. S. (1962). *The structure of scientific revolutions*. Chicago: The University of Chicago Press.
10. Kuhn, Th. S. (1974,1977). Objectivity, value judgment, and theory choice. In Th. S. Kuhn (Ed.), *The essential tension – Selected studies in scientific tradition and change* (pp. 320-339). Chicago: The University of Chicago Press.
11. Morreau, M. (2014). Mr. Fit, Mr. Simplicity and Mr. Scope: From social choice to theory choice. *Erkenntnis*, *79*, 1253-1286.
12. Morreau, M. (2015). Theory choice and social choice: Kuhn vindicated. *Mind, 124*(*493*), 239-262.
13. Okasha, S. (2011). Theory choice and social choice: Kuhn vs. Arrow. *Mind*, *129* (*477*), 83-115.
14. Rawls, J. (1971). *A theory of justice*. Cambridge, MA.: Harvard University Press.
15. Rizza, D. (2014). Arrow's theorem and theory choice. *Synthese, 191,* 1847-1856.
16. Saari, D. G. (1995). *Geometry of voting*. Heidelberg, Berlin: Springer Verlag.
17. Samuelson, P. A. (1967). Arrow's mathematical politics. In S. Hook (Ed.), *Human values and economic policy* (pp. 41-52). New York: New York University Press.
18. Scanlon, Th. M. (1998). *What we owe to each other*. Cambridge, MA.: Harvard University Press.
19. Scott, M. J., & Antonsson, E. K. (1999). Arrow's theorem and engineering design decision making. *Research in Engineering Design*, *11*, 218-228.
20. Sen, A. (1970). *Collective choice and social welfare*. San Francisco: Holden-Day.

21. Sen, A. (1977). On weights and measures: Informational constraints in social welfare analysis. *Econometrica, 45,* 1539-1572.
22. Sen, A. (2009). *The idea of justice*. Cambridge, MA.: Harvard University Press.
23. Stegenga, J. (2015). Theory choice and social choice: Okasha versus Sen. *Mind, 124(493),* 263-277.
24. Strasnick, St. (1976). Social choice and the derivation of Rawls's difference principle. *The Journal of Philosophy, 73,* 85-90.
25. Suppes, P. (1966). Some formal models of grading principles. *Synthese, 16,* 284-306.
26. Svensson, L.-G. (1977). *Social justice and fair distributions*. Lund: Lund University.
27. Temkin, L. S. (1993). *Inequality*. Oxford: Oxford University Press.
28. Weber, M. (2011). Experimentation versus theory choice: A social-epistemological approach". In H. B. Schmidt, D. Sirtes, & M. Weber (Eds.), *Collective epistemology*, (pp. 1-26). Frankfurt: Ontos Verlag

Alternatives	Power (bhp)	Weight (kg) criterion 1	Power/weight (bhp/t) criterion 2	Cost Criterion 3
a	300	1000	300	6300
b	330	1050	314	6600
c	350	1100	318	7000
d	400	1300	308	6200

Table 1: A multi-criteria decision problem

The three criteria “weight”, “power/weight” and “cost”, taken together, cannot be arranged in a single-peaked fashion.