Aggregating Judgment in Scientific Practice

May 8, 2015

Abstract

The endeavor to apply the theory of judgment aggregation (TJA) to scientific practice has two main motivating factors. First, this application proceeds from the hope that this new theory could provide a new perspective on well-known problems of the philosophy of science. Secondly, it stems from the desire to extend this theory to new domains, and even to seemingly intractable cases. A peculiar feature of scientific social practices is that they connect people that belong to diverse generations. This paper argues that standard theories of judgment aggregation cannot apply to scientific practice, since science is a temporally extended process that involves both different individuals and different hypotheses during that process. Thus, for example, we seem to have no idea how to determine the judgments of dead scientists about theoretical alternatives that were proposed after their death. It then proposes an algorithm for judgment aggregation to try to address some of these challenges.

Keywords

Judgment aggregation theory, scientific practice, intergenerational practice.

1 the ideal of an intergenerational aggregation procedure

1.1 The challenge of extending TJA to scientific practice

Broadly understood, a theory of aggregation seeks to discover how a group could produce judgments or decisions that outperform the cognitive abilities

of the individuals composing this group. The normative reflection on aggregation procedures could stem from an empirical observation as well as from a thought experiment: depending on the aggregation or voting rules that are accepted, the same set of individual judgments or preferences can generate various collective preferences or decisions. Therefore, an aggregation theory is a normative theory that ought to compare the ability of several possible procedures to generate results that can meet the goals that motivate the implementation of this procedure.

A scientific practice can be roughly characterized as a collective endeavor aiming at finding true explanatory principles in a given area. I will refer to the judgments on these explanatory principles as "theoretical judgments". TJA seems like a perfect framework within which to investigate this practice, since one can hope to aggregate various judgments on explanatory principles that are discussed within a group.

At first glance, the TJA seems to easily model and evaluate the functioning of scientific communities. It has recently shown that sociological accounts of these communities, for instance the description by Knorr-Cetina (1999) of the HEP experiments, fit List's following theorem:

> For any group size n (divisible by k), there exists an individual (positive and negative) reliability level $r^* > r$ such that the following holds: if, by specializing on some proposition p, individuals achieve a reliability above $r*$ on p, then the majority judgment on p in a subgroup of $\frac{n}{k}$ specialists (each with reliability r^* on p) is more reliable than the majority judgment on p in the original group of non-specialists (each with reliability r on p). (List, 2005)

However, science is not only a sociological phenomenon, but also a historical phenomenon: it is not only a process where living people interact, but also an intergenerational process of cooperation. There are three kinds of interactions within this practice: interactions among the living, interactions between the living and the deceased, and interactions between living and virtual people, that is the future generations. As long as this second aspect of scientific social practice, that is, the intergenerational component, is not accounted for, the application of TJA to the scientific communities only constitutes a partial extension of this theory to the scientific practice.

These intergenerational interactions can be observed in various social and historical phenomena:

1. **Transmission**: this phenomenon brings about knowledge cumulatively, across several dimensions: knowledge about singular facts, knowledge on phenomenal regularities within phenomena, technical knowledge (both material and formal).

2. **Collaboration**: Some questions cannot be raised without other questions having been previously raised and, often, answered. I will call these questions that are raised by a given generation "intragenerational questions".

3. **Epistemic dependence**: Some questions cannot be raised as long as an adequate vocabulary has not been established.

4. **Disagreement**: the formulation of new hypotheses that break away from past hypotheses as an answer to what I will refer to as "transgenerational questions". For instance, on the same question, "Does void exist?", Aristotle opposes Democritus, Galileo opposes Aristotle, Torricelli opposes Galileo, and so on.

Since this last aspect of the intergenerational relationship involves judgments, there are rational hopes that it could be modeled within the TJA framework. The goal of this paper is to attempt to do so and to answer the following question: can there exist an aggregation procedure of intergenerational judgments that generates more reliable theoretical judgments than those which would be produced by each generation working in isolation?

1.2 The challenge of the intergenerational feature of scientific practice

Intergenerational disagreement calls into question several aspects of the standard view about science and thereby raises at least four important problems of the philosophy of science. The historical continuity/discontinuity problem deals with the threat of incommensurability that follows from the existence of theoretical breaks between successive generations. The problem of unconceived alternatives, recently formulated by Stanford (2006), is a new version of the underdetermination thesis: reflecting on the fact that the history of science offers several cases where new theoretical hypotheses have been formulated that had not been conceived before, Stanford maintains that nothing can guarantee that our best present theories will not be replaced sooner or later by empirically equivalent but logically incompatible theories. This claim is used to argue that we should not be optimistic

concerning the ability of scientific practice to provide us with correct theories. The existential problem stems from what Kierkegaard (2009) calls "the anticipation of the retrospection": the efforts of researchers seems absurd, meaningless, or aimless, if the result of their findings can be totally reversed by the subsequent, unpredictable and unconceivable interventions of future generations. The scientific agency problem consists in concluding, in light of contradicting scientific opinions over time, that that the set of scientists could be regarded as a collective subject.

It is rational to hope that TJA could offer a framework to solve these problems by helping us to decide between a revisionist and a traditional conception of science. The revisionist perspective denies to science properties that are traditionally associated with it: historical continuity, epistemic optimism, existential satisfaction, the unity of the enterprise. As revisionists focuses on the empirical fact of intergenerational disagreement, the normative reflection associated with TJA could help to establish that their claims are not well-grounded: TJA could constructively prove that the intergenerational character of the scientific practice does not necessarily lead to renouncing the traditional view of science.

Indeed, if a procedure of aggregation of intergenerational judgments (PAIJ) could be constructed, the following alternative theses could be drawn. Against incommensurability, the successive opinions could be regarded as the elements of an homogeneous collective judgment. Against epistemic pessimism, a more optimistic conception of science could be sustained where the ability of the present theory to access truth is not cancelled by the possibility of unconceived alternatives. Against existential despair, a scientist could legitimately believe that his efforts will contribute to the collective judgment. Against agency nominalism, we could see scientists as struggling together towards the same goal.

2 The obstacle of the interviewer's temporal mobility

An aggregation procedure is a group of rules that permits us to derive a collective judgment from the empirical basis of the procedure. It has its starting point in the set of judgments that a group of individuals holds on one or several propositions. I will call "empirical basis of the procedure" the set of information concerning these judgments. The empirical basis of the procedure must be collected from individual judgments found in a social reality and then be represented. I will call "interviewer" the ideal individual

who performs these two tasks. The classical cases of TJA are situations where these individuals (in our case, scientists) are living contemporaries, so that the procedure can be started after a phase of judgment collection: as these two steps are clearly separated, the temporal neutrality of the aggregation procedure is guaranteed: the decision to launch the procedure at one moment rather than at another does not affect its result.

The main obstacle to a PAIJ is the fact that depending on the historical location of the interviewer, the aspect of the empirical base varies. In what follows, I am going first to distinguish the different species of these variations and the specific kind of obstacle they present to the work of the ideal interviewer. The consequence of this variability of the empirical basis is the possibility that the application of the same set of rules to the same reality produces opposite results. These results are paradoxes, understood as contrary propositions that are nonetheless equally justified. The confrontation of these distortions with the temporal neutrality ideal will be used to define the limits and the rules that a PAIJ should accept. This will give me a ground on which to first elaborate a technique of representation of the empirical basis of the procedure.

2.1 Problems stemming from the variation of the empirical basis

The first two problems come from the fact that past and present scientists do not have an opinion about hypotheses of which they cannot conceive, whereas present and future generations can formulate judgments on hypotheses belonging to the past. I will refer to this phenomenon as the intergenerational epistemic asymmetry (IEA).

- **The spiritism problem**: Given IEA, how can we obtain from the deceased to take a stand on theoretical possibilities that they ignored?
- **The oracle problem**: given IEA, it is not only impossible for us to know what will be the hypotheses that will be accepted tomorrow, but also to foresee what will be the stance of future generations on propositions that we accept today.
- **The generational individuation problem**: the concept of generations, understood as referring to groups, is not a natural species: human individuals, as living creatures, are continuously born. It is therefore difficult to determine which individuals to include in a gen-

eration. As an example, we could just as well count within a century only one generation or millions of them.

- **The problem of the fixation of origin**: The "interviewer" can either take into account every judgments recorded since the formulation of a question, or only those judgments that are subsequent to a given date. It he wanted to, the interviewer could justify such a disqualification on the ground that before a certain historical moment, people did not have the technical or social means to deal rationally with a question.

- **The *kairos* problem**: the undertaking of starting an aggregation procedure is the pragmatical equivalent to a declaration designating the present moment as the one starting from which opinions do not count anymore for the collective opinion. The aggregation action is necessarily egocentric: the content it offers is not objective, but is structured according to the temporal position of the interviewer. Future generations are disqualified.

- Since a period is delimited by two frontiers, the fixation of origine and the *kairos* problems form together what could be called the **periodization problem**.

2.2 Paradoxes that result from these problems

I now intend to prove that the periodization problem enables the interviewer to reverse any result of any procedure. The representation of the empirical basis of the procedure appears to be subjective in two ways: it not only necessarily rests on an arbitrary choice, but it also necessarily involves self-reference. I refer to this result as an impossibility theorem: the representation of the empirical basis cannot be objective. It follows from this impossibility theorem that the ideal interviewer is endowed with ultimate authority on the matter of what the collective judgment should be: whatever the procedure used is, he is the one whose decision determines what the collective judgment should be.

2.2.1 a) The majority procedure

This procedure regards the judgment that the most people take on a proposition as the group's judgment. There exist several possible historical situations relevant to that procedure. Either there are only two possible answers

to the transgenerational question or there are more than two. In the first case, generations that share the same judgment can follow one another or can alternate with generations that do not share their judgment. Here, I am going to focus on a simple situation, since the same pattern of reasoning can easily be applied to complex situations. Let us therefore suppose a real situation where there are two possible continuous opinions: either the majority opinion includes the present one, or it excludes it. In the first case, it is possible to turn the majority opinion into a minority opinion by deciding to dissociate the moment of aggregation and the posterior limit of the period. In the second case, to turn the majority opinion into the minority, it is only necessary to move the anterior limit of the period.

2.2.2 b) The unanimity procedure

This procedure generates a collective judgment from a set of judgments only if this judgment is unanimous within a given group. Two relevant initial situations are possible: unanimity or non-unanimity. If there is no unanimity, it is always possible to generate unanimity by narrowing the empirical basis of the procedure, whereas if there is unanimity, it is always possible to destroy it by broadening the empirical basis, in the direction either of the past or the future.

2.2.3 The dictatorial procedure

This procedure derives the collective judgment from the judgment of one member of the group. The philosophy of science has opposed two kinds of dictatorial procedure: critical rationalism and traditionalism. In the first case, the dictatorship is assumed by the present generation, whereas in the second case, it is ascribed to a definite past generation.

2.3 Restriction on the constitution of the empirical basis

My goal here is to avoid paradoxes and to escape the impossibility theorem by questioning the assumptions of the theorem and accepting a set of restrictions that should be followed by the interviewer. In the next section, I will offer a procedure that respects these restrictions. For now, I would like to show how to formalize the empirical basis of the procedure.

Let us start with the list of the three restrictions:

1. No overlapping generations. (This restriction enables the interviewer to count generations.)

2. No oracle.

3. No spiritism.

As a consequence, an aggregation procedure is bound to be egocentric.

These restrictions set the basis for a set of interview techniques designed to take them into account. The subjective effects of this egocentric point of view can be circumscribed by a "metaprocedure comparison": the intergenerationnal situation render possible a succession of procedures and, therefore, the comparison of their results. Even though each one is egocentric, the outcome of the comparison can be objective.

The "non overlapping" restriction enables the interviewer to use a chart which represents the judgments of each generation on a set of possible answers to a question. I will note as G_q the generation that has formulated a transgenerational question and as G_p the present generation. I will note as G_u the ultimate generation, that is the generation that will have the correct answer and know all possible hypotheses. From the "no oracle" and "no spiritism" conditions follow the need of specific notation: I will use Y for a positive judgment, N for a negative judgment, $\emptyset$ for the ignorance of a proposition,and ? for the cases when the interviewer does not know what a generation answers. With this notational system, the interviewer could represent the empirical basis of a procedure with such a chart as the one that follows:

	G_q	$G_{(p+1)}$	G_p	$G...$	G_p	$G_{(p+1)}$	...	G_u
H_1	$\emptyset$	$\emptyset$	?	...	?	N	?	?
H_2	Y	N	Y	...	Y	Y	?	?
H_3	$\emptyset$	Y	N	...	N	N	?	?

Table 1: Empirical basis chart

3 Description of an intergenerational judgment aggregation procedure

3.1 Definition of the goals of the procedure

This procedure can have two objectives: knowledge and progressiveness.

3.1.1 a) Definition of knowledge

List & Puppe (2009) suggests that to evaluate the ability of a procedure to meet the knowledge challenge, the concept of knowledge offered by Nozick could be useful. This definition enables List to apply the same concept of knowledge to both individuals and groups; to make possible quantitative evaluations of the relationship between an agent and knowledge; and finally to build the concepts of positive and negative competence. According to this definition, one can say that an individual S or a group G knows that p if and only if the four following conditions are met:

1. S or$G$ believes or assert that p.
2. p is true
3. If p had not been true, S or G would not have believed or asserted that p.
4. There are no cases where p could have been true without S or G believing or asserting it.

Drawing on this definition, List uses conditional probabilities to build the concepts of positive and negative competence. The positive competence of an individual S is the probability that S asserts p, given that p is true. The negative competence of S is the probability that S rejects p, given that p is false. The advantage of these notions is that they enable us to quantify the trustworthiness of groups or individuals in order to decide whether the aggregation procedure has been useful or to compare the efficiency of two procedures.

3.1.2 b) Definition of progressiveness

Progressiveness cannot be the goal of only one procedure but of several. Indeed, this goal cannot be pursued unless we compare the results of several successive procedures. The succession of several procedures is progressive if the fact of successively implementing these procedures at moments t_1 and t_2 guarantees the increase of our knowledge. Formulated in terms of competence: there is a progression if, between t_1 and t_2, the positive and negative competence have increased. Thus:

- There has been a progression between t_1 and t_2 if $P_{t_1}(B_p \wedge p) < P_{t_2}(B_p \wedge p)$.

- There has been a progression between tt_1 and t_2 if $P_{t_1}(B_p \wedge p) > P_{t_2}(B_p \wedge p)$.

3.2 Elucidation of the presuppositions of the procedure

The general conditions of a procedure must be compared to the description of the scientific practice.

3.2.1 a) Independence of generations

It is necessary to clearly see the difference between information dependence and judgment dependence. The present generation depends on past generations in order to know what hypotheses have been conceived and what judgments have been made of them. However, this does not imply that the present judgment has to be the same as the past one. Indeed, past non overlapping generations have no power over the present generation. Moreover, the experimental method, broadly understood, supposes that each individual, and moreover each generation, has the right, if not the obligation, to criticize hypotheses of the past by confronting them with his own experiments.

3.2.2 b) The dissymmetry of positive and negative competences

I have defined the scientific practice as the research, in a given area, of the true explicative principles, that is, propositions from which a set of phenomena can be deduced. As it is possible to deduce true consequences from false propositions, several propositions exist which could fulfill the explanative demand concerning the phenomena. As a consequence of this abductive nature of scientific research, several logically incompatible but empirically equivalent theories exist throughout history. And the more there are possible hypotheses, and the more choosing one of these hypotheses is risky.

This problem is deepened by the fact that we cannot know how many hypotheses are possible, because of the conjecture concerning unconceived alternatives. Therefore, it seems that, if hypothesis h is true, the chance that a generation believes h is extremely low: partly because of their own creativity, the positive reliability of scientists is much lower that 50%. On the other hand, the experimental method seems to guarantee a high level of negative competence to scientists, since their ability to recognize that a hypothesis is false does not vary with the number of possible hypotheses. Since the historical and empirical knowledge of new generations is steadily increased

by the empirical results produced by the previous generation, the negative competence of a generation is always higher than the negative competence of any previous generation.

3.3 Description of the procedure

Within the theoretical framework that I have offered it should be obvious that neither the dictatorial nor the unanimity procedure could meet the knowledge challenge. Since the positive competence is lower than 0.5 %, the Condorcet Theorem excludes the majority procedure. For the same reason, a dictatorial procedure would have a very low positive competence, since this competence would be equal to the competence of an isolated generation. The dictatorial procedure would still produce better results than the majority procedure.

The first rule of the procedure is to derive a positive judgment of the group if this judgment is unanimous between the formulation of the intergenerational question and the moment of the procedure.

If there is no such unanimity, the second rule comprises two steps. The first step of the procedure consists in listing the possible known answers to the transgenerational question. The group holds a negative judgment on a proposition if since this proposition has been rejected, this judgment has not been reversed.

The second step of the procedure consists in considering all the negative judgments selected in the first step as the premise of an eliminative reasoning which has two premises: the first one is the conjunction of all possible and known hypotheses; the second one is the conjunction of the premises that are rejected. The proposition or propositions that have not been rejected are equated with the collective judgment of the scientific intergenerational community at the moment of the procedure.

3.4 Evaluation of the epistemic results of this procedure

What is the probability that the proposition on which there is an unanimous positive judgment is correct? Let us start with some hypotheses. There is a higher probability that individuals that are biased toward truth reach an agreement than if they are biased towards error, provided that there is only one way of being right and several ways of being wrong. We can therefore suppose that

$$P(C|R) > P(C|W) \wedge P(C|R) > P(C) \tag{1}$$

with: $P(C)$, the probability of a consensus; $P(C/R)$ the probability of a consensus knowing that scientists are right; $P(C/W)$ the probability of a consensus knowing that scientists are wrong; $P(R)$, the probability that a scientist is correct is identical to his positive competence. Using Bayes theorem, we can submit that:

$$P(R\ |C) = \frac{P(C|R)(P(R)}{P(C)} \tag{2}$$

Assuming that the more generations are considered, the lower the probability of a consensus is, we can conclude that the more generations are involved in an unanimous decision, the more probable it is that these generations support a correct judgment. The first rule of the procedure is therefore progressive, since between two moments, the number of generations increases so that the probability of knowledge raised.

What is the epistemic outcome of the application of the second rule? Suppose this time that the negative competence of scientists is high. Suppose also that the number of possible answers to a transgenerational question is limited even though it is unknown. According to a simple reasoning, the more hypotheses we have eliminated, the likelier it is that the remaining hypothesis is the correct one. Therefore, the more a procedure eliminates hypotheses, the likelier it is to yield a correct answer. As this number increases with the number of generations, the progressiveness challenge is satisfied.

Conclusion

The result of this twofold procedure seems, at first sight, strange. It looks like a dictatorial procedure since the collective judgment is the same as the judgment of the last generation. What then is the interest of an intergenerational aggregation procedure? It shows that, provided that several generations face the same question and provided that they have the right to judge independently, the probability that the last generation holds a true judgment is much higher than the positive reliability of this last generation by itself.

This result shows that PAIJ could help to answer the problems of the philosophy of science introduced in the first section. This procedure answers the existential problems since it does not discard the efforts of a given generation, but rather takes them into account. It allows us to regard intergenerational disagreement as a collaborative organization aiming at producing

an eliminative reasoning. It offers a solution to the question of how to articulate continuity and discontinuity within the history of science, in order to avoid the most radical versions of the incommensurability thesis. And, last but not least, it enables us to consider the possibility of unconceived alternatives without being driven to despair.