The Scylla and Charybdis of Analytical Sociology

N. Olsson-Yaouzis

June 2, 2013

1 Introduction

Philosophers of science have recently turned their attentions towards the role mechanisms play in explanations. [See, e.g., Bunge [1997], Machamer et al. [2000]] The increased interest has, in part, been fueled by dissatisfaction with the traditional law-based accounts of explanation.

In particular, the mechanism accounts of explanations have been well received in the philosophy of the social sciences. This has, in part, to do with the fact that rational choice sociologists, such as Jon Elster [1983], as early as the 1980s suggested that explanations requires mechanisms.

Recently, proponents of the so-called *analytical sociology* movement have defended mechanism based explanations in the social sciences. For example, Peter Hedström [2005, p. 2] begins a book on the principles of analytical sociology by arguing that mechanism based explanations are the most appropriate type of explanation for the social sciences. Elsewhere, Peter Hedström and Petri Ylikoski [2010, p. 58] argue that the only way to satisfy analytical sociology's demand for precision and clarity is to make explanations explicit by providing articulated mechanisms.

Many proponents of the mechanism account also believe that the call for mechanisms can settle the dispute between methodological individualism and holism in favor of some sort of individualism. For example, Hedström and Ylikoski [2010, p. 59] claim that a basic point of the mechanism perspective is that explanations that simply relate macro properties to each other are unsatisfactory.

Although the mechanism account can solve some of the problems that have plagued the traditional law-based account of explanation, there are reasons for being skeptical about the claim that it can be used to rule out methodological holism. Therefore, it seems as analytical sociology should not restrict its search for mechanisms described in terms of individual properties.

In this essay we will argue that analytical sociology face a problem when they hold that proper explanations in the social sciences should consist of mechanisms in terms of individuals and their properties. In order to pull this off, they have to navigate between the Scylla of full-fledged reductionism to physics, and the Charybdis of holism where explanations in terms of social facts has to be accepted.

In section 2 we will present and motivate the mechanism account of explanation. We will also offer examples of a mechanism that are judged appropriate by analytical sociology. In section 3 we will look at analytical sociology's reasons for blocking the reduction from the level of individual and their properties, to a lower level of explanation. Next, in section 4, we will look at the reasons for motivating the reduction from the macro-social level to the individual level. We will argue that proponents of analytical sociology fail to show that explanations in terms of macro-social properties are inappropriate. The reason is that the reduction from the social level can be blocked for the same reason as the reduction from the individual level is blocked. Finally, in section 5, we wll conclude and point out the sound intuition behind the demand for mechanisms in terms of indivudal properties.

2 Mechanisms and analytical sociology

The best way of introducing the mechanism account of explanation is by describing some of the problems of the covering law account.

According to the covering law account an event is explained if a statement describing the event, E, can be showed to follow from some relevant initial conditions, I, and law-like generalizations, L.[1] The initial conditions and generalizations are called the *explanans*, and the description of the event to be explained is called the *explanandum*. The covering law account can be illustrated with the following simple example:

1. Whenever air pressure increases, the mercury in the barometer rises.
2. Air pressure increases.
3. Therefore, the mercury in the barometer rises.

Here we have a deductive argument where the explanandum, 3, is the conclusion of the explanans, 1 and 2. This kind of argument can be called a deductive-nomological (DN) argument.

[1] See, e.g., C.G. Hempel [1966].

The covering law account of explanation has at least two appealing features. First it seems to capture some of our intuitive understanding of what it means to explain a phenomenon: if the law-like generalization and the initial conditions are true, then the explanandum must occur.

The second somewhat appealing feature is that the covering law account does not entail any metaphysical commitments concerning causality. The law like generalization can be thought of as describing a Humean constant conjunction: whenever an event of a type A occurs, then (shortly afterwards) an event of type B occurs. As such, the generalizations do not commit us to a non-reducible concept of causality.

Although the second feature might appeal to philosopher with metaphysical commitment problems, it is also responsible for some of counter-intuitive implications. The first problem is that many DN-arguments with true premises do not appear to be explanatory. Wesley Salmon [2006, p. 50] illustrates this point with the following example:

1. Butch takes birth control pills.
2. Butch is a man.
3. No man who takes birth control pills becomes pregnant.
4. Therefore, Butch does not become pregnant.

Although the explanandum, 4, can be deduced from the true statements in the explanans 1, 2, and 3, it does not seem as a proper explanation of why Butch does not become pregnant. The problem, according to Salmon, is that the premises are explanatory irrelevant to the conclusion.

A second problem for the covering law model is that it fails to account for the asymmetry of explanation. This problem was raised by Sylvain Bromberger [1966] who pointed out that if we can construct a valid DN-argument with the position of the sun, the height of the Empire State Building, and the laws of optics as premises and the conclusion that the building's shadow is x meters; then it is possible to use the same statements to construct a valid DN-argument where the height of the Empire State Building is the conclusion and the length of the shadow part of the explanans. Contrary to what the covering law account implies, it seems as only the first DN-argument is explanatory. Intuitively, it is the height of the building that explains the length of the shadow and not the other way around.[2]

[2]For a contrary opinion see Bas van van Fraassen [1980, pp. 132-5].

Finally, James Woodward [2003, p. 183] has been pointed out that although law-like generalizations play a central role in the covering law account, no satisfactory criteria for distinguishing genuine laws from accidental generalizations have been produced. The latter type of generalizations may be true, but they are not true in virtue of any laws of nature. For example, the generalization "all stones in this box are white" may be true, but it does not seem to be the kind of generalization that can be used in a scientific explanation. Some have concluded that, until such criteria are produced, we should abandon not only the covering law model but all law-based accounts of explanation.

2.1 The mechanism account

Hedström and Ylikoski [2010, p. 55] points to the above three problems as reasons for abandoning the covering law account of explanation. They also add the observation that there are very few covering laws, especially in the special sciences such as biology, psychology and the social sciences. The last observation is problematic in so far as we believe that the special sciences have produced satisfactory explanations.

One early formulation of the mechanism account can be found in Elster [1983, p. 24] where he claim that "to explain is to provide a mechanism, to open up the black box and show the nuts and bolts, the cogs and wheels of the internal machinery." Although there is something intuitively appealing with the demand that an explanation should "open up the black box" and uncover the underlying "cogs and wheels," it is obviously formulated in metaphorical terms and therefore in need of further explication. Hedström and Ylikoski [2010, p. 50] somewhat improve on Elster's formulation by claiming that "proper explanation should detail the cogs and wheels of the causal process through which the outcome to be explained was brought about."

In order to produce criteria that allows us to identify a mechanism explanation we need to get rid of the metaphors in the above accounts. This will require a more precise description of what mechanism is. Fortunately, Hedström [2005, p. 25] provide this. According to him, a social mechanism is defined as "a constellation of entities and activities that are organized such that they regularly bring about a particular type of outcome." He goes on to claim that "we explain an observed phenomena by referring to the social mechanism by which such phenomena are regularly brought about."

A mechanism explanation seems to capture some of the same intuitions as a covering law explanation. Just as a covering law explanations shows

that we have no reason to be surprised that the explanandum occurred given the provided explanans, so does a mechanism explanation show us that the explanandum was to be expected as the outcome of a mechanism that regularly brings about this kind of events. After all, once we have identified that the entities and activities that regularly bring about a certain type of event are present, then much of our surprise should disappear.

The main difference between the covering law and the mechanism accounts of explanations is that the mechanism account is committed to, what Hedström calls, causal realism. That is, unlike the covering law account that is compatible with a regularity interpretation of the generalizations, the mechanism account is committed to the existence of causation and causal processes. For example, on the mechanism account it is not the observation of a constant conjunction of storm and barometer readings that allows us to explain that the barometer changes when the storm approaches, rather it is the existence of some causal process that links the approach of the storm to the barometer change that allows us to explain the barometer change.

For a philosopher with metaphysical commitment problems this might seem as a cost. At the same time, it is the causal realism that allows the mechanism account to avoid the problems of the covering law account.

For example, the lack of a causal link between Butch's consumption of birth control pills and him not getting pregnant is the reason why the premises are explanatory irrelevant with respect to him not getting pregnant. The mechanism account, unlike the covering law account, can use real causal processes to account for explanatory relevance.

Similarly, the mechanism account can use causality to account for the asymmetry of explanation. Causal processes have a direction, and if a causes b, then it cannot be the case that b causes a. Thus, an explanation of the length of the shadow requires a mechanism that describes the causal process that leads from the sun via the Empire State Building to its shadow. Since the causal chain goes from the sun to the shadow, we cannot use the same mechanism and backtrack and explain the height of the building in terms of the length of the shadow.

Thus, it can be argued that the mechanism account's ability to avoid these problems more than enough makes up for its increased metaphysical commitments. Let us, therefore, accept Hedström's mechanism account of explanation and use an example to illustrate what he believes this implies for the methodology of the social sciences.

We can illustrate what seems to be the methodological implications with the help of a thought-example from Glenn Loury [2002, p. 30]. Assume that we want to explain our observation that young black men find it much

harder to get a taxi than young non-black men. One explanation in terms of structural discrimination might be that there are racist structures in the observed society that induce this type of behavior in the taxi drivers.[3] This explanation is a typical "holistic" explanation as it relates a racist *structure* to the difficulties of the young black male population to get a cab. Note that if we accept that structures and populations are entities that can engage in activities, then there should not be any problem of formulating this explanation in terms of a macro-mechanisms. However, according to Hedström "an important thrust of the analytical approach is that *actors and actions* are the entities and activities of the mechanisms explaining [social phenomena]." [Hedström, 2005, p. 28, emphasis added]

Fortunately, Loury offers an alternative explanation of the phenomenon that is more in line with what Hedström's approach. According to [Loury, 2002, p. 30], the inability to get a cab can be explained in terms of, what he calls, self-confirming stereotypes. Assume that initially we have two groups of prospective young male cab riders, black and non-black, and that each group have the same share of criminal elements. That is, if 2% of the prospective black taxi riders are robbers, then 2% of the non-blacks are robbers as well. Now, assume that some of the cab drivers have a false belief that there are more robbers in the black population than in the non-black population. Being afraid of getting robbed, these taxi drivers will refrain from picking up young black men. This, in turn, will make it slightly more difficult for young black men to get a taxi. This will increase their taxi-search cost, either in terms of the time they have to wait to get picked up or in terms of humiliation of not getting a cab. For some of the young black male the increased search cost will make them opt for public transport instead of taxi. Since is is likely that the cut off point for switching to other means of transport is higher for robbers than non-robbers, this will increase the proportion of robbers in the black cab riding population as compared to the non-black cab riding population. This, in turn, will cause cab drivers to form a correct belief that there is a higher risk of getting robbed by young black male cab riders, then by young non-black male cab riders. And so the process goes on until the vast majority of young black male cab riders are actually robbers (although there are proportionally as many robbers in the non-black population as in the black population).

In Loury's explanation there are entities in the form of taxi drivers and cab riders. They have certain properties. For example, the cab drivers have

[3] Admittedly, this is probably a caricature of what a proponent of a structural discrimination account would say. But it will suffice for our purposes.

beliefs about risks of getting robbed, and the cab riders have beliefs about search costs; they have desires to not get robbed, and to minimize waiting time. The entities engage in activities like picking up prospective cab riders, or waiting for cabs. Finally, these activities and properties form a causal chain that leads up to the phenomenon that we are interested in explaining.

As we have mentioned above, according to analytical sociology it is the actors and actions that are the core entities and activities of the mechanisms explaining social phenomena. Now, the challenge for analytical sociology is to motivate the reduction from the social level to the individual level, while at the same time resisting the reduction from the individual level to the physical level. We will begin by investigating whether analytical sociology can avoid its Scylla in the form of full-fledged reduction, and then move on to see whether it can avoid its Charybdis of holism.

3 Avoiding Scylla by blocking the reduction

According to Salmon's [1984] mechanism account of explanations in the natural sciences, an adequate explanation requires that we formulate the fundamental causal mechanism leading up the phenomenon we are interested in.[4] For example, assume that we are interested in explaining the arc a search light traces across the sky. The movement of the light beam is, on Salmon's account, a pseudo-process since earlier positions of the beam does not cause the later positions. Our explanation should therefore ignore this pseudo-process, and instead look for the true causal process. The true causal process takes place in the generation of the light itself, and the movement of the lamp. According to Salmon, an adequate explanation would require us to find the fundamental causal process.

Salmon's requirement is deeply reductionistic. Stathis Psillos [2002, p. 282] points out, for example, that on Salmon's account the law of ideal gases is merely a "lawful regularity" since it does not display the underlying causal mechanism that connects the macroscopic parameters (i.e., temperature, pressure, and volume). In the case of gases, the causal law is provided by the molecular-kinetic theory of gases. Furthermore, on Salmon's account, it is only genuine causal laws that are explanatory, and whatever explanatory import "lawful regularities" may have, is parasitic on the fact that the regularity can itself be explained by the causal law.

Salmon's causal fundamentalism provides a somewhat unambiguous answer to the question of what entities a mechanism consists of: mechanisms

[4]See also Mayes [2005].

consist of whatever entities are found at the bottommost level. Furthermore, if we search for and find Salmon-type mechanisms then we can be confident that we have opened all glaring "black boxes," and have laid bare all "cogs and wheels."

However, if adequate explanation requires a specification of the fundamental causal process, and if the fundamental causal process is to be spelled out in the language of physics, then all other sciences are in trouble. Salmon's account seems to entail that there are much fewer explanations than we might have believed. It does, for example, seem as what we thought were explanations in sociology, biology, even chemistry, are not genuine explanations. Furthermore, since we lack a sociological analogue to the molecular-kinetic theory of gases, we might even have to conclude that our suggested sociological explanations lack explanatory import even in the parasitic sense.

Since Salmon has no particular interest in the special sciences, this might be a bullet he is willing to bite. For proponents of analytical sociology, on the other hand, this conclusion should be unacceptable. The question then becomes how to block the reduction to what Salmon takes to be fundamental causal level.

3.1 Multiple realizability

The traditional method of blocking reduction to physics while maintaining physicalist intuitions, is to accept what Jerry Fodor [1994, p. 689] calls "token physicalism" while denying "type physicalism."

Token physicalism is the claim that all *events* the sciences talk about are physical events. That is, each event mentioned by chemistry, biology, or sociology is a physical event. Token physicalism prevents the sciences from talking about events that cannot be given a physical description such as (a naive interpretation of) the progression of the Hegelian Geist or demonic possession. In Loury's taxi driver example, token physicalism implies that each mental event is also a physical event. For example, *this* taxi driver's *present* belief that young black men are more likely to be robbers is identical to some physical event.

Type physicalism, on the other hand, is the stronger claim that each *property* mentioned by the laws of any science is a physical property. According to this doctrine all properties mentioned by chemistry, biology and sociology are physical properties. According to type physicalism any property that is not identical to a physical property should not be used in the sciences. Applied to Loury's taxi driver example, type physicalism implies

that the mental properties are identical to some physical properties. For example, the property of being a belief that young black men are more likely to be robbers is identical to some physical properties.

According to Fodor and others, token physicalism is acceptable whereas type physicalism is not. The reason is that because mental properties are multiply realized by physical properties. To see the problem let us look at a causal explanation of why a taxi driver does not stop for a young black man:

> the belief that stopping for young black men increases the risk of getting robbed, and the desire to not get robbed causes the taxi driver to not stop for a young black men.

In order to see why it is not possible to reduce this explanation to a physical explanation, let us accept the following claims:

1. each instance of an action can be given a purely physical description (e.g., as pure behavior),

2. each instance of a mental state (e.g., desires and beliefs) can be given physical description (e.g., as brain states), and

Even if we accept these claims it does not seem as we can translate our explanation in terms of mental properties to an explanation in terms of physical properties. The reason is that each mental state type and action type is multiply realized by physical properties. For example, the action of not stopping for a young black man can be realized by many different behavioral descriptions. At one time this action can be described as the man sitting in the driver's seat keeps his foot steady on the gas pedal, at a second time he slightly increases the foot's pressure, at a third time he holds the steering wheel with both hands, at a fourth with one hand, etc. In order to give a full physical description of the action, we have to offer a physical description of every instance of the action. Not only every instance in the actual world, but also every possible instance that would qualify as not stopping for young black men would need to be given a physical description. Thus, the physical description of the action will be, as Fodor puts it, wildly disjunctive:

> x is the action of not stopping for young black men if, and only if, x is the set of bodily movements 1, or x is the set of bodily movements 2, ..., or x is the set of bodily movements n.

Although n is probably a large number, let us concede that it does not approach infinity. Thus, it should, in principle, be possible to give a complete physical description of the action of not stopping for young black men.

There is no reason to think that it is easier to offer physical descriptions of the taxi driver's mental states. For example, note that the exact propositional content of a desire can differ while still being characterized as a desire to not get robbed. Both the desire that *it should not be the case that I get held up at gunpoint* and the desire that *it should not be the case that I get held up with a knife* seem to be desires to not get robbed. Since the propositional content differs, it should be the case that different brain states can realize a desire not get robbed. There is no reason to think that beliefs are less multiply realizable than desires.

Thus, physical descriptions of desires and beliefs will be as (wildly) disjunctive as physical descriptions of the actions:

> x is the belief that stopping for young black men increases the risk being robbed if, and only if, x is the set of brain states 1, or x is the set of brain states 2, ..., or x is the set of brain states m.
>
> x is the desire to not get robbed if, and only if, x is the set of brain states 1*, or x is the set of brain states 2*, ..., or x is the set of brain states l*.

If we bring the physical descriptions together we can offer the following physical explanation of the taxi driver not stopping for young black men:

> the set of brain states 1, the set of brain states 2, ..., the set of brain states m and the set of brain states 1*, or the set of brain states 2*, ..., or the set of brain states l* caused the set of bodily movements 1, or the set of bodily movements 2, ..., the set of bodily movements n.

The problem with the physical explanation is not that it is incorrect. After all, it is just a redescription of the mental explanation. So if the explanation in terms of mental states is correct, then so is the physical explanation. Rather the problem is, according to Fodor, that the language of physics fail to carve the mental world at its joints. The wildly disjunctive physical description of mental states and actions show that physics is not fit to describe this level of reality.

Even if multiply realizability does not block inter-theoretic reduction there are good reasons for preferring the explanation in terms of mental

states to the physical explanations. One reason is that the disjunctive physical explanations are unappealing when compared to the much simpler explanations in terms of mental states. How exactly are we to go about identifying the explanans and explanandum in the disjunctive physical explanation? Although there are plenty of problems involved in identifying beliefs, desires, and actions, these problems are much smaller than the problems of using our current technology to identify exact descriptions of brain states or set of bodily movements. Considering the problems of testing, confirming, and using the disjunctive explanation, explanations of actions in the language of physics are, to say the least, impractical.

3.2 The relevance reason

However, impracticality is not the main reason offered by Hedström for resisting the reduction from the mental to the physical. According to Hedström [2005, p. 27], the main problem of reducing mental explanations is that physical explanations are unlikely to be of much *relevance* to sociology. The idea is developed in Hedström and Ylikoski [2010, p. 52] where it is claimed that

> The why or how questions one is addressing determines what the representation of the mechanism should include in order to be explanatory. Only by knowing the nature of the explanatory task at hand can one determine which details of a mechanism are relevant to include and the appropriate degree of abstraction.

So even if mental properties were not multiply realized by physical properties, the analytical sociologist could resist the reduction to physics by an appeal to relevance. In order to see what relevance amounts to let us say something more about why-questions.

Although many philosophers have treated explanations as answers to why-questions[5], it is probably Bas van Fraassen [1980] who has made the most extensive use of this observation. According to van Fraassen an explanation is an answer to a why-question of the type "why P (rather than X)?" P is called the topic and X the contrast class of the question.

The inclusion of a contrast class X is a necessary part of a fully specified why-questions. To see this consider the question "why did the taxi driver drive on?" This question can be interpreted as either "why did the *taxi driver* drive on (rather than the bus driver)?" or "why did the taxi driver

[5]See, e.g., Hempel [1965] and Bromberger [1966]

drive on (rather than stop for an ice cream)?" It is only be specifying the contrast class X that we can distinguish between the two interpretations.

According to van Fraassen the contrast class, X, is usually implicitly determined by the context the question is asked in. If asked by sociologists, then the first contrast class may be implicitly included, whereas if asked by marketers, the second contrast class may be implicitly included.

van Fraassen goes on to argue that in order to fully specify the why-question we need to specify what kind of question we are interested in. He does this by introducing a *relevance relation*, R, that has to hold between the topic-contrast class pair, and the answer, A, in order for A to be part of an appropriate answer.

van Fraassen includes the relevance relation in order to specify the standards for an appropriate answer in a given context. Think of the question "Why did the match catch fire rather than fizzle?" and three possible answers: because...

1. A_1 = ... it was struck.

2. A_2 = ... the striking surface was made of sand, powdered glass, and a chemical called "red phosphorus." The head of a match was made of sulphur, glass powder, and an oxidizing agent. When the match was struck on the striking surface of its box, the friction caused by the glass powder rubbing together produced enough heat to turn a very small amount of the red phosphorus into white phosphorus, which catches fire in air. This amount of heat was enough to start the chemical reaction that used the oxidizing agent to produce oxygen gas. The heat and oxygen gas caused the sulphur to burst into flame which then caused the wood of the match to catch fire.[6]

In the context of a child posing the question to a parent, then the relevance relation would pick A_1 as the appropriate answer. When posed to an engineer, A_2 could be the most appropriate answer. Thus, we need the relevance relation in order to identify whether the why-question is part of, e.g., everyday discourse or scientific inquiry. This example indicates that also the relevance relation is usually determined by the context the question is asked.

Thus, two scientific disciplines can ask, what appears to be, the same why-question while requesting different answers. The reason is that when a scientific why-question is asked, the scientific context will determine the

[6]This excellent answer was provided by the Science Theatre at Michigan State University, http://www.pa.msu.edu/sciencet/ask_st/092596.html, accessed on 13 May 2011.

relevant answers. Thus, when the questions "why does the majority of fatal airplane accidents take place at landing (rather at some other time)?" is posed to an engineer, the relevant answer will be in terms of mechanical failure. When asked to a psychologists, the relevant answer might be in terms of human failures.

Let us call this the *relevance reason* for blocking the reduction to the physical level. Let us also repeat that this will be a reason for resisting the reduction that does not depend on whether social or mental mechanisms are multiply realized by physical mechanisms.

Before moving on let us also point out that both the multiple realizability and relevance reasons allow the proponent of the mechanism account to remain committed to mechanisms as "an irreducible causal notion"[7] and also accept that genuine causal processes are only found at the fundamental physical level. After all, even if causality obtains only between events on the fundamental physical level these events do not have to be described in the language of fundamental physics. Since explanations use descriptions of events rather than the events themselves, and since we can offer a near infinite number of descriptions of a causal process leading from an event to another, we can offer a near infinite number of explanations of the same event.

That is, one explanation for each description of the events involved in the process. After all, if an event is explained under one description, then it will also be explained under a (correct) redescription. For example, if it is true that the sinking of the Titanic was caused by a collision with an iceberg, then it is also true that the sinking of the fastest ship on the Atlantic 1912 was caused by a collision with an iceberg.[8]

Similarly, if 1) it is true that brain state i caused bodily movements j; and 2) if having brain state i refers to the same event as having the belief that stopping for young black men increases the risk of getting robbed and the desire to not get robbed; and 3) if bodily movements j refer to the same event as not stopping for a young black man; then 4) it is also true that the belief that stopping for young black men increases the risk of getting robbed, and the desire to not get robbed caused the taxi driver to not stop for a young black men.

Thus, it seems possible to avoid Scylla by making a distinction between causes and explanations in terms of causes. Thus, even if the causal processes take place on the fundamental physical level, the fact that psycholog-

[7] E.g., Hedström and Ylikoski [2010, p. 50] and Ylikoski [2012, p. 3].

[8] Given that the Titanic was also the fastest ship on the Atlantic 1912.

ical and social types are multiple realized by physical types and that social and psychological descriptions are more relevant to sociology than physical descriptions gives us good reasons to prefer explanations in psychological or social terms.

4 Avoiding Charybdis by motivating the (other) reduction

Analytical sociology does, however, not only have to avoid the Scylla of reduction, but also the Charybdis of holism. The problem is that the same reasons that have been provided to argue that mechanisms described in terms of individuals and their properties are permitted, can be used to argue that mechanisms in terms of social entities and their properties are methodologically kosher.

In order to motivate the reduction from the macro-level to the individual level, the analytical sociologist need to show why these arguments are unavailable to the macro-sociologist. We will begin with the argument from multiple realizability and then turn to the relevance reason.

4.1 Holism and multiple realizability

Both Hedström [2005, pp. 70-74] and Ylikoski [2012, p. 8] discuss multiple realizability in connection with critical realism that holds that structures have autonomous ontological existence and causal powers. One problem with critical realism, according to Hedström, is that structures are unobservables. Therefore, if the critical realists want to argue that structures have autonomous causal powers, then they have to provide a method that allows us to reliably identify these structures. Hedström holds that until such a method has been provided, critical realists are unjustified in treating structures as having autonomous causal powers. This seems to be a reasonable objection.

Furthermore, Ylikoski's discussion of the argument from multiple realizability focuses on its inability to support the claim that structures have autonomous causal powers. This should not come as a surprise since the argument from multiple realizability is used to show that higher order *predicates* are not be reduced to lower order *predicates*. Above we showed that there are good reasons to believe that mental types are multiply realized by brain state types, and that therefore it is difficult to translate our descriptions in terms of mental types into descriptions in terms of brain state types.

Since we have accepted that genuine causality can only be found on the bottommost level, this argument is not meant to show that mental states have autonomous causal powers. It is only meant to show that we have reasons to prefer the description of the causal process in terms of mental states rather than brain states (or, for that matter, elementary particles).

Similarly, the fact that social types are multiply realized by individual types does not show that social structures have autonomous causal powers. Hedström suggests that the argument from multiple realizability could be used to support methodological holism. Methodological holism would be the view that although all genuine (social) causal processes involve non-holistic entities and their properties, we should describe these processes in terms of holistic properties. The fact that holistic properties are multiply realized by, e.g., individual properties provides an argument for methodological holism in the same way as it provided an argument for resisting the reduction from individual properties to physical properties.

However, Hedström [2005, p. 74] goes on to point out that methodological holism entails that

> from a causal point of view, a correlation between two social phenomena will therefore always be epiphenomenal and, in this sense, spurious.

Well, this is in a sense true. But as we have pointed out this should come as no surprise for a honest methodological holist. For the methodological holist it does not matter whether genuine causal powers rest in individuals or elementary particles, all that matters is that holistic properties cannot or should not be reduced to individual properties.

Unless Hedström wants to argue that individuals are privileged in the sense that have autonomous causal powers, the methodological holist is in the very same position as the analytical sociologist who wants do defend mechanisms in terms of individuals against reduction to mechanisms in terms of physical properties. In other words, the argument from multiple realizability seems to be available for analytical sociologists and methodological holists alike.

Hedström [2005, p. 73] points out that he doubts that social properties are multiply realized by individual properties. However, since there seems to be a near infinite number of ways the social predicates "... is money" or "... is a firm" can be realized by physical and individual properties common sense seems to speak in favor for this thesis. Thus, the burden of proof should here rest on the analytical sociologist who wants to deny methodological holists the argument for multiple realizabilty.

4.2 Holism and relevance

However, it might be possible for the analytical sociologist to deny the relevance of the argument from multiple realizability by arguing, e.g., that multiple realizability does not block reduction. This might be the reason why neither Hedström nor Ylikoski uses this argument to block the reduction to physical explanations.

The problem is, of course, that methodological holists can use the relevance argument to argue that in some sociological contexts, the relevant description of the mechanism is in terms of macro-properties. For example, it might be claimed that some sociological why-questions are best answered in terms of xenophobia rather than people's beliefs and desires. It is difficult to see how the analytical sociologist can deny this argument to the methodological individualists without being forced to give it up themselves.

In other words, the macro-sociologist can use the same arguments to block the reduction from the macro to the individual level as the analytical sociologist used to block the reduction from the individual to the physical level. Thus, in order to motivate the reduction from the macro to the individual level, some additional argument is needed.

4.3 Reduction to avoid mistaking correlation for causation

Hedström [2005, p. 29] provides another reason for preferring mechanisms described in terms of individuals to mechanisms in terms of holistic properties. It is that it will allow us to reduce the risk of mistaking spurious correlations for genuine causation.

Two events are spuriously correlated when there is a third event causing both. For example, we may discover that growing up with many bookshelves and attending college are correlated and infer that having bookshelves cause college attendance. These would, however, be spuriously correlated if there is a third event, such as having highly educated parents, that causes both having many bookshelves and college attendance. If we correctly identified the causal mechanism leading up to college attendance then this would allow us to rule out that having bookshelves cause college attendance.

Or as Hedström and Ylikoski [2010, p. 54] put it: "The knowledge that there is a mechanism through which X influences Y supports the inference that X is a cause of Y." However, as we have argued, there are many ways to describe a genuine causal mechanism. One may be in terms of physical properties, a second in terms of individual properties, and a third in terms of holistic properties. If all refer to the same causal mechanism through

which X influence Y, then all should be able to support the inference that X is a cause of Y. Thus, the demand that we should have knowledge of the mechanism does not speak against explanations in terms of holistic properties.

However, the problem for the scientist who focuses solely on holistic properties might be epistemic. Elster [1985, p. 5] argues that in order to rule out spurious correlations and, thus, to justify our causal claims we have to make the temporal and spatial gap between explanans and explanandum as small as possible. Thus, correlation between holistic properties does not provide sufficient reason to believe that a suggested macro-process accurately describes a genuine causal mechanism. In the bookshelf-college example it might be argued that once we describe the mechanism in terms of the individuals and their desires and beliefs we will discover that there is no causal connection between growing up with bookshelves and attending college.

However, the problem is that focusing on individuals and their properties does not eliminate the risk of mistaking spurious correlation for causation. It is easy to imagine situations where two events described in terms of individuals are strongly correlated while being caused by some third event. For example, we might discover a strong correlation between having anti-social attitudes and criminal actions, and conclude that having anti-social attitudes cause criminal actions. However, unbeknownst to us there might be a gene that cause both anti-social attitudes and criminal actions. In order to avoid confusing a spurious correlation with causality we would have to descend to the biological level. Unfortunately, the search for mechanisms described in terms of biology does not completely eliminate the risk of confusing correlation for causality. The risk will be minimized (and perhaps eliminated) first when we reach the fundamental physical level.

At this point the dispute between the analytical sociologist and methodological holist boils down to a tradeoff between relevance and risk reduction. The methodological holist would hold that the relevance gains outweigh the risk of confusing correlation with causation. The analytical sociologist, on the other hand, would have to deny this and show that the gains of reducing the risk by going from the social to the individual level outweighs the loss of relevance. In order to avoid the Charybdis of having to accept holostic explanations, the analytical sociologist has to appeal to some normative reasons for preferring some degree of risk reduction to some relevance increase.

5 Concluding remarks: the sound intuition of analytical sociology

It seems as analytical sociology can avoid Scylla of reduction and Charybdis of holism. However, the problem was that the strongest arguments (multiple realizability and relevance) for avoiding reduction of sociology to physics, could also be used to block the reduction from macro-sociology to individualism.

Thus, Charybdis of holism could not be avoided with the help of a metaphysical argument showing that any descriptions on the macro-social level is reducible to a description on the individual level. Nor could it be avoided by an appeal to relevance.

In the end, Charybdis was avoided by appealing to the cost of mistaking correlation for causation. However, in order to use this as a reason to prefer mechanism explanations in terms of individuals, the analytical sociologist had to say something about the tradeoff between the gains of reducing the risk of confusing correlation for causation and the cost incurred by not being able to offer relevant answers. Whatever they will say they have to be careful not to put too much weight on risk reduction since this may force us to search for mechanisms described in sub-individualistic terms.

Finally, let us admit that there seems to be something intuitively appealing in the analytical sociologists' demand for individualistic mechanisms. The intutive appeal seem to rest on the idea that we should have some idea of how our explanations connect to each other. That is, there is something deeply unsatisfactory with an explanation that seems to be completely disconnected from the rest of the sciences. Harold Kincaid [1994] illustrates this intuition with two holistic entities: Adam Smith's *Invisible hand* and Hegel's *Weltgeist*. According to Kincaid, Smith's explanation in terms of the Invisible hand are appropriate since he gives us some account of how it is connected to the rest of our sciences. Explanations in terms of Hegel's Weltgeist seem to be unsatisfactory in part because no account of how it connects to the other sciences is given.

However, it is far from obvious that this intuition can only be satisfied by a ful specification of how a holistic mechanism is realized by a mechanism on a lower level. Think, for example, of the causal claim that Bengt's French improved because he spent some time in France last summer. We accept this as an explanation of why Bengt's French improved although nobody has access to the exact physical processes that connect Bengt spending some time in France with the improvement of his French. It is enough for us

to know that it seems reasonable that there exists a process in a person's brain such that when he spends time in a Francophone country his French improves.

Similarly, the methodological holist can accept that if there is no plausible way for a holistic mechanism to be connected to an individualistic mechanism, then it ought to be rejected. However, from this it does not follow that each holistic mechanism must be accompanied by a fully specified individualistic mechanism. It might be sufficient that the holist provides a rough sketch of the lower level mechanism in order to justify the holistic mechanism.

References

S. Bromberger. Why-questions. In R. Colodny, editor, *Mind and Cosmos: Essays in Contemporary Science and Philosophy*. Pittsburgh: University of Pittsburgh Press, 1966.

M. Bunge. Mechanism and explanation. *Philosophy of the Social Sciences*, 27(4):410 – 465, 1997.

J. Elster. *Explaining Technical Change: A Case Study in the Philosophy of Science*. Cambridge: Cambridge University Press, 1983.

J. Elster. *Making Sense of Marx*. Cambridge: Cambridge University Press, 1985.

J.A. Fodor. Special sciences (or: The disunity of science as a working hypothesis). In M. Martin and L.C. McIntyre, editors, *Readings in the Philosophy of the Social Sciences*, pages 687–699. Cambridge, MA: MIT Press, 1994.

P. Hedström. *Dissecting the Social*. Cambridge: Cambridge University Press, 2005.

P. Hedström and P. Ylikoski. Causal mechanisms in the social sciences. *The Annual Review of Sociology*, 36:49–67, 2010.

C.G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press, 1965.

C.G. Hempel. *The Philosophy of Natural Science*. Prentice-Hall, 1966.

H. Kincaid. Reduction, explanation, and individualism. In M. Martin and L.C. McIntyre, editors, *Readings in the Philosophy of the Social Sciences*, pages 497–513. Cambridge, MA: MIT Press, 1994.

G.C. Loury. *The Anatomy of Racial Inequality*. Cambridge, Mass.: Harvard University Press, 2002.

P. Machamer, L. Darden, and C.F. Craver. Thinking about mechanisms. *Philosophy of Science*, 67(1):1–25, 2000.

R.G. Mayes. Theories of explanation. In R. Bishop, editor, *Internet Encyclopedia of Philosophy*. 2005. URL `http://www.iep.utm.edu/explanat/`.

S Psillos. *Causation Explanations*. Acumen Publishing, 2002.

W. Salmon. *Scientific explanation and the causal structure of the world*. Princeton, N.J.: Princeton University Press, 1984.

W.C. Salmon. *Four Decades of Scientific Explanation*. Pittsburgh: University of Pittsburgh Press, 2006.

B.C. van Fraassen. *The Scientific Image*. Oxford: Oxford University Press, 1980.

J. Woodward. *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press, 2003.

P. Ylikoski. Micro, macro, and mechanisms. In H. Kincaid, editor, *The Oxford Handbook of Philosophy of Social Sciences*, pages 21–45. Oxford: Oxford University Press, 2012.