

Thinking with the Coleman Boat

Petri Ylikoski, Linköping University and University of Helsinki

Thinking with the Coleman Boat

Petri Ylikoski

Institute for Analytical Sociology

Linköping University

Department of Social Research

University of Helsinki

1. Introduction

The Coleman diagram is one of the most famous theoretical diagrams in sociology. It is often referred to as Coleman's boat, and in the German-speaking world it is sometimes called Coleman's bathtub. In this paper I will simply call it "the diagram." James Coleman introduced the diagram in his 1986 paper, and used it extensively in Coleman (1987) and his *magnum opus*, *Foundations of Social Theory* (1990). The diagram did not catch on immediately, but especially after Hedström and Swedberg used a modified version of it in their 1998 anthology about social mechanisms, it has become quite a popular representation of the micro-macro challenge in sociology. It is particularly prominent in the context of analytical sociology (Hedström & Ylikoski 2011), but is also used by many sociologists who are not associated with this movement. Apart from sociology, the diagram has also been utilized in debates in economics (Janssen 1993; Hartley 1997) and political science (Beach & Pedersen 2013; Rohlfing 2012).

The diagram is interesting mainly for two reasons. First, it addresses one of the core issues in social theory: how to relate micro and macro. Sociologists are interested in explaining the emergence,

persistence, and change of large-scale social events, properties, and processes. Their relation to smaller-scale social events and processes (like individual behavior, social interactions, etc.) provides a thorny set of problems. Sociologists want to understand both how large-scale things (the macro) influence the smaller-scale things (the micro) and how the macro is comprised of micro-scale events and activities. The diagram provides a systematic way to think about these micro-macro relations. Sociologists are increasingly treating the micro-macro link as a substantial theoretical challenge in their empirical research, rather than an issue of competition between sociological approaches or as an abstract philosophical concern. This makes the right understanding of the diagram a pertinent issue in philosophy of the social sciences.

Second, the diagram is a very rare visual representation in sociological theory. While most diagrams drawn by sociological theorists are mere summaries of verbal argument, or are otherwise rather primitive compared to the visual representations developed in other sciences (Lynch 1991; Turner 2010), the Coleman diagram provides an example of representation that can serve as a cognitive tool for sociological thinking. It is “a genuine tool of production,” as C. Wright Mills (1959: 213) puts it. The diagram provides a systematic scheme for articulating social explanations and their presuppositions. While the content of the scheme could in principle be verbally articulated, the easily remembered diagram provides a more cognitively efficient representation. The abstract form of the diagram makes it easy to adapt to various applications. It raises a sequence of questions that a sociologist must answer to be able to claim to have theoretically understood the social phenomenon. The productiveness of the diagram is based on the ability of these questions to trigger new lines of thinking about the phenomenon. The diagram is therefore a productive cognitive tool, rather than a mere device for summarizing what is already known.

However, an abstract diagram is easy to misunderstand. Even a cursory look at various interpretations, modifications, and applications of the diagram demonstrates that there is no general

agreement about what the diagram represents. For example, many critics have taken it as a summary of Coleman's theoretical ideas. Thus, it has been interpreted as a statement of the implications of methodological individualism, rational choice theory, or more generally, reductionism. It has also been interpreted as an argument for the causal impotence of macro properties. I will argue that all these interpretations are false. The diagram is instead a cognitive tool that is independent of Coleman's other ideas in sociological theory. Thus while Coleman claimed to be a methodological individualist (at least in some sense), and he employed rational choice theory, the usability of the diagram is independent of his views on these issues. In other words, people who are not methodological individualists or advocates of the rational choice theory can also use the diagram. Similarly, I will show what is wrong with the common practice of interpreting the diagram using concepts from the philosophy of mind.

The aim of this article is to provide a rational reconstruction of the diagram as a tool for social scientific theorizing. I will show how Coleman uses the diagram and how it can be employed as a tool for theoretical thinking. I will also demonstrate how it can be used to clarify the nature of the micro-macro challenge in social explanation. The structure of the article is as following. Section 2 will describe the diagram and its elements. Section 3 will show various ways in which Coleman used the diagram in his work. Section 4 discusses more systematically issues related to the interpretation of the diagram, and Section 5 will provide a diagnosis of some recent interpretations of the diagram. Finally, Section 6 focuses on some important limitations of the diagram.

2. The nodes and arrows

To facilitate the references to the diagram I have named the nodes with capital letters and the arrows with numbers, as shown in Figure 1. This representation differs slightly from Coleman's original diagram. Coleman himself did not use letters to name the nodes, as in his diagrams the nodes

always contained a reference to the relevant causal factor or variable. Similarly, the numbering of the arrows is not present in Coleman's original diagrams, but the use of these numbers has become standard practice although the numbering varies between authors. Finally, and more significantly, arrow 4 is not present in all of Coleman's own diagrams. Yet as the following discussion will show, the occasional absence of arrows has no theoretical significance: in the use of the diagram arrow 4 plays no systematic role. Thus the fact that Coleman included the arrow in the diagram in his 1987 and 1990 publications, but not in those of 1986 and 1992, does not reflect any changes in his thinking.

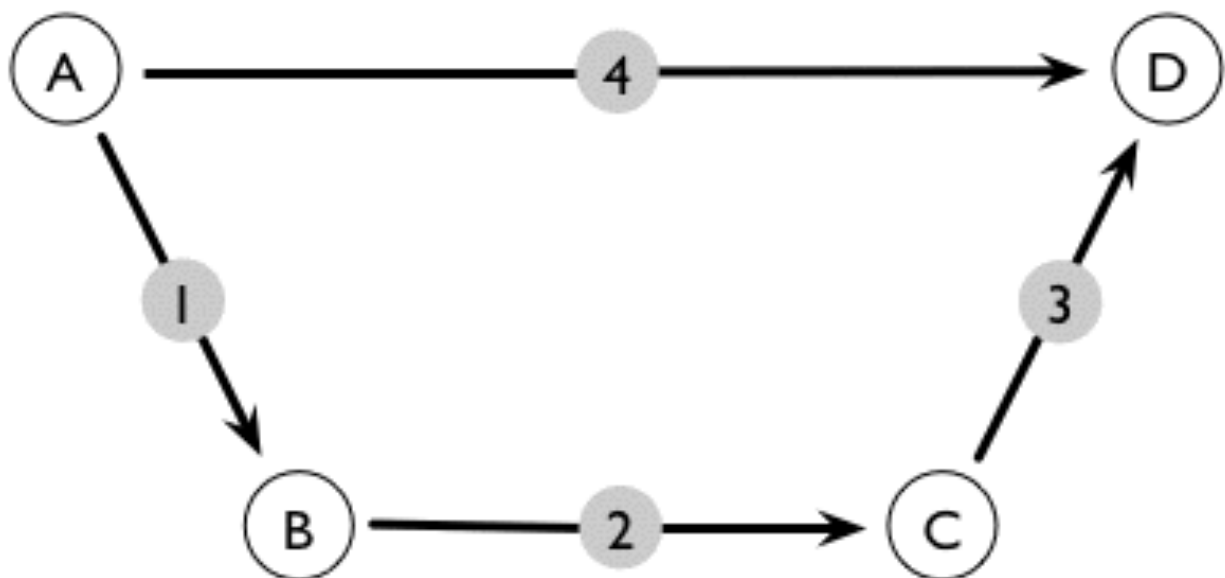


Figure 1. The Coleman diagram

This paper will focus on Coleman's use of the diagram. However, it is worth noting that Coleman did not invent the diagram out of nothing. One clear precursor can be found in David McClennan's book *The Achieving Society* (1961: 47). McClennan's diagram has the same recognizable arrows, and he also uses the same example, Weber's thesis in *Protestant Ethic*. According to Siegwart Lindeberg, Coleman learned about McClennan's diagram from Lindeberg in the early 1980s, who himself had learned about it from McClennan's lectures at Harvard in the 1960s (Barbera 2007: 44).

It is not clear whether McClennan and Lindeberg used the diagram in exactly the same way as Coleman. Be this as it may, it seems that all the later-published uses of the diagram refer to Coleman (or to authors who refer to Coleman), which justifies my focus on Coleman.¹ Without Coleman, the diagram might have been completely forgotten. The diagram has also been modified, changed, or reinterpreted in many ways by later authors (see, for example, Hedström & Svedberg 1996; Hedström & Swedberg 1998; Abell 2000; Udéhn 2001; Manzo 2007; Abell, Felin & Foss 2008; Hagan & Raymond-Richmond 2009; Sampson 2012; Thornton, Ocasio & Lounsbury 2012; Beach & Pedersen 2013; Erikson 2014). However, apart from Hedström and Swedberg's 1998 mechanistic interpretation, these modifications have not been adopted for more general use. Due to limitations of space, I will not discuss these alternative interpretations in this paper.

Understanding the diagram presupposes a proper understanding of the nodes and arrows. Nodes A and D refer to various sorts of macro social facts that figure as *explanantia* and *explananda* in sociological explanations. Basically, A's are extra-individual social factors that might be cited as causes of social phenomena and that might influence individuals. D's are macro social facts to be explained. These facts can be quite diverse. Among Coleman's own sociological *explananda* are revolutions, economic structures, inequalities, panics, election outcomes, and the disintegration of families. Coleman's notion of macro is flexible: its scale can vary from two-person social interaction to organizations and nation states (Coleman 1986, 1987, 1990). In addition, the A-facts can be quite different in their scale. Thus he says that A-facts could be "... characteristics of their social environments, ranging from family to friends to larger social contexts" (1990: 1). The facts that can fill node A are not privileged: in principle, all A's can (and should) serve as D's for other explanations. In other words, there is no assumption that there are unexplainable social facts. It is crucial to note that A and D do not refer to the whole "macro-level," but only to particular facts

¹ For a discussion of precursors of the Coleman diagram, see Raub and Voss (2016). Among the various precursors I find Hernes (1976) especially interesting and close to Coleman's ideas.

relevant for the hypothesis under consideration. As I will argue in detail later, the Coleman diagram is quite different from the supervenience/realization diagrams used in the philosophical mental causation debate.

Node B refers to the properties of agents and their situations. In most of Coleman's own examples the B-facts are about values, opportunities or interests, but this set can be extended also to beliefs, desires, goals, values, preferences, motives, emotions, habits (or habitus), routines, scripts, heuristics, cognitive schema, and identities (social and psychological), as all these notions play a similar role in sociological explanation. They are all theoretical concepts used in the explanation of human behavior, and their role is to mediate between social influences and individual behavior. Some of these notions might ultimately be too vague to be useful in the social sciences, but the evaluation of these concepts is not of concern here. In this context, it is important to understand their role in the explanation of action. Sociologists are not primarily interested in explaining B-facts. The B-facts are important because they explain behavior (C-facts) and mediate the influence of A-facts on it. So, while there might be theoretical reasons to avoid some of these notions, the explanation of action requires that some concepts like these be employed.

Node C refers to behavioral outcomes. Depending on the context, they could be choices, behaviors, or actions. These individual outcomes are important *explananda* for sociology, but they are also a step towards explaining macro outcomes (D-facts). Coleman allows that the B- and C- facts do not necessarily concern individual persons. They could also "be institutions within the system or subgroups" (1990: 2). They could also be corporate agents. Therefore, not only is Coleman's notion of *macro* flexible, so is his notion of *micro*. This implies that the diagram cannot be understood as an argument for methodological individualism (contra Jepperson & Meyer 2011). What is important for Coleman is the agency, not individuality (or personhood) of agents (Ylikoski 2016). Thus he writes that "for some investigations, corporate bodies such as formal organizations are usefully

regarded as purposive actors, though in other research and theory in sociology, the coherence of their action would itself be taken as problematic” (1986: 1312).

The arrows are a bit more difficult to interpret. They are sometimes misunderstood as representing relations of reduction or logical derivation. However, the best way to understand them is to take them as explanatory relations that are based on causal dependencies. It is clear that Coleman intends that arrows 1, 2, and 3 are to be interpreted causally. For example, he articulates the arrows in the Weber example in the following way (1990: 8):

- * Arrow 1: “Protestant religious doctrine generates certain values in its adherents.”
- * Arrow 2 “Individuals with certain values adopt certain kinds of orientations to economic behavior.”
- * Arrow 3 “Certain orientations to economic behavior on the part of individuals help bring about capitalist economic organization in a society.”

All of these are straightforward causal claims. “Generation” and “bringing about” are causal notions, and it is also obvious that in the second sentence the values are assumed to causally influence the economic activities. I will later discuss some problems with this straightforward causal interpretation, but for the current purposes it is sufficient to accept that the arrows reflect explanatory dependencies. The more detailed interpretation of these dependencies is contingent on one’s account explanation. Thus, we have a Dutch interpretation that relies on a covering-law account of explanation and reconstructs the diagram in terms of deductive arguments (see Lindenberg 1981; Raub, Buskens, & van Assen 2011 for details). The alternative Swedish interpretation (Hedström & Swedberg 1998) interprets the arrows in terms of causal mechanisms. While both interpretations might be legitimate, I think they are too demanding for many uses of the diagram. It is more helpful to interpret the arrows in terms of a less ambitious account of

explanation. Thus, although it is nice to have a mechanism-based understanding of the explanatory dependencies represented in the diagram, the applicability of the diagram should not depend on having such knowledge.

My suggestion is that the arrows be interpreted in terms of difference-making. Thus in the case of arrow 1, for example, some change in the A-variable brings about a change in the B-variable. This is a kind of minimal explanatory dependency that is easy to comprehend in terms of (imagined) causal interventions: changing A to A* would bring about B changing to B*. For example, a government changes the legislation on parental leave (intervention in the A-variable) which brings about changes in workers' opportunities to take care of their small children (change in the B-variable) which in turn brings about changes in their annual working time (C-variable). This idea is fully compatible with mechanisms-based thinking: one can always ask by which mechanisms were these changes brought about. However, one might not have that information available, or multiple causal pathways (mechanisms) may be at work that might be in opposition to each other, so one would still like to know whether there is some sort of net effect on B. The difference-making interpretation also provides a good way to conceptualize the structural presuppositions of macro-micro and micro-macro claims: they are background conditions – or a causal field – that make the foreground causal claims possible.

Finally, there is arrow 4, which is sometimes absent from Coleman's own diagrams. It is the starting point of the analysis that employs the diagram. It might be a correlation between A- and D-variables, or it could be a hypothetical causal connection between these variables. The purpose of the diagram is to help in determining whether the relation between A and D is truly causal, and if it is, why it holds. In this way the use of the diagram begins with dissatisfaction with the A-D link. Further, two concerns motivate this dissatisfaction.

The first concern is related to causation. Contrary to some suggestions (Opp 2011), Coleman is not against the legitimacy of macro causal claims in general. However, they do raise the problem of justification. Quite often there is not enough macro data (too few cases to observe) to fully support these causal claims, and it is very rare that experimental evidence is possible. One way to make these causal claims more secure is to provide a set of mechanisms by which the suggested cause brings about the effect. The function of the diagram is to help in figuring out whether such mechanisms can exist. For example, if a university believes that changes in professors' salaries increases the scientific quality of research produced by the university as a whole, it is a good idea to articulate how the supposed causal chains are assumed to work. If one cannot provide a plausible and empirically supported story, there is every reason to be skeptical about the claimed causal dependence. Furthermore, if such a causal narrative can be provided, it might be possible to test the individual elements of the chain (the arrows 1, 2, 3). This would provide empirical support for the claim, even if it impossible to directly test the macro claim.

The second concern is related to sociological explanation. Here the basic idea is that even if it might be possible to secure the causal relation between these two macro variables, this would still be theoretically insufficient. One would also want to know why the dependence holds, in other words, one would want to know the underlying mechanism. This explanatory ambition is shared by all theoretically oriented sciences, so the concern is in no way particular to sociology. The ability to answer the how-questions behind the simple causal claim expands one's understanding (Ylikoski & Kuorikoski 2010). For example, the information about the mechanisms helps to understand under which background conditions the causal dependency holds, how it would change when some of the background conditions change, and what can be done to prevent or sustain the causal outcome (Ylikoski 2011).

3. The uses of the diagram

Now that we have a basic understanding of the components of the diagram, it is time to look at the things that can be done with it. I will start by considering how Coleman uses the diagram in his papers.

First, the diagram is a tool for thinking about particular sociological explanations. It can be used to *illustrate* particular explanatory ideas – for example, the marriage "market" and other matching processes (1987: 159-160), and problems with collective decision-making (1987: 169-171, 1990: 400-402) – or to *compare* competing explanations, for example the explanations for the panic in a burning theater (1987: 161-163, see also 1990: 203-209). However, the diagram is not merely an illustration tool. Coleman uses it to *point out missing crucial elements* in proposed explanations. For example, this is the main point of his discussion of Weber's thesis in *Protestant Ethic* (1987: 154-155, 1990: 6-10). He argues that "some sort of combined or joint or aggregate effect of the economic behavior of many individuals in bringing about capitalist development is being proposed. It is here, however, that Weber's analysis is almost totally silent" (1990: 9). The frustration theories of revolution (1987:156-157, 1990: 478-479) receive a similar treatment: "The third relation is implicit, a simple aggregation of individual aggression to produce a social product, that is, a revolution. Yet a revolution involves organization and the interplay of actions on the part of a number of actors." (1990: 10) Both examples share the same basic failure: "the micro-to-macro transition is made simply aggregation of individual orientations, attitudes, or beliefs. If, however, the theoretical problem is one of involving the functioning of a social system [...] then it should be obvious that the appropriate transition cannot involve the simple aggregation of individual behavior" (1990: 10). The contribution of the diagram is to point out which parts of explanatory narrative are missing. Finally, Coleman uses the diagram also to *articulate the structural assumptions of explanations*. A good example of this is his discussion of the influence of educational inequality on income (1987: 163-166) and the impact of training programs on youth

employment (1987: 166-168, 1990: 642-645). In both cases, the diagram is used to trigger articulations of the suggested explanations' presuppositions so that their credibility can be evaluated.

At a more general level, the diagram provides *an account of the explanatory tasks of sociology*. It gives an ideal description of the full sociological explanatory narrative. Rather than just citing a relation between two macro variables, an ideal sociological explanation provides an account of how the suggested cause brings about a change in individuals' (or other agents') beliefs, desires, or opportunities, how these changes in turn bring about changes in their behavior, and, finally, shows how these behavioral changes bring about the macro change to be explained. Not only does the diagram point to these three analytically distinct explanatory subtasks, it also gives an account of how they are related. This scheme makes it possible to point out areas of theory that are underdeveloped in sociology. The main point of the above-quoted criticism of Weber and the frustration theories of revolution was to show that sociologists have not paid enough attention to the micro-macro links in their theories. Only when these links are adequately considered is the sociological understanding of a phenomenon theoretically satisfactory.

The diagram can also be used in thinking about foundational issues in social theory. First, the diagram is useful in highlighting the non-trivial nature of the problem of microfoundations. The explanations of C-facts and D-facts involve structural assumptions that are crucial for understanding the explanatory dependencies. This is Coleman's key message: the micro-macro relations are only aggregative in special cases. Second, the diagram can be used to illustrate the role of "the theory of action" in sociological explanation. While pure BC-explanations fall largely outside the scope of sociology, both ABC- and BCD-explanations presuppose BC-processes, which implies that assumptions about psychological processes play an important role in sociological explanations. Finally, the diagram also presents the basic idea of mechanism-based explanation

already explained above: sociological explanations that merely connect two macro variables (AD-explanations) are not theoretically satisfactory and need to be supplemented by an account of the micro process underlying them. The point of mechanism-based explanation is to show how the causal (or non-causal) relation between A and D is brought about. This helps us both to justify (or criticize) the suggested causal relation, and to have an explanatory understanding of it.

In this context it is also worth observing what the diagram is not. First, contrary to Jepperson & Meyer (2011), the diagram is not intended to be a serious contribution to Weber scholarship. Coleman uses Weber theory in his *Protestant Ethic* as an example because it is widely known. He does not make broad claims about the scope of Weber's more general theory, but simply points out an obvious missing piece of the popular interpretation of the explanatory narrative presented in Weber's 1905 book. The popular interpretation might be wrong, but this does not diminish its value as an illustrative example. Second, while Coleman claims that he is a methodological individualist, the fact that he uses the diagram to articulate structural presuppositions of explanations should raise an alarm among those (Jepperson & Meyer 2011; Little 2012) who think that the diagram *presupposes* some form of methodological individualism. When one understands how the diagram is actually used, it is impossible to think that arrow 1 is an arrow of reductive explanation. What about the relation between the diagram and the rational choice theory? Coleman does use the diagram to illustrate his ideas inspired by rational choice theory – especially in *Foundations of Social Theory* – but it should also be clear from the way the diagram works that rational choice theory is not the only theory of action that can be used to deal with arrow 2. This implies that there is no essential connection to rational choice theory. The diagram is a tool for thinking, not a metaphysical or theoretical statement.

4. Thinking the diagram through

In this section I take a closer look at elements of the diagram and discuss some difficulties related to their interpretation. I argue that while interpreting the diagram is more complicated than usually assumed, the ambiguities and problems do not pose a fundamental challenge to the diagram's usefulness in analyzing sociological explanations.

4.1. Macro to micro

Causal influences from macro to micro are a central concern in sociology. One could even say that they are a defining feature of a distinctly sociological approach to the social world. This is apparent when we consider the alternative diagram, shown in Figure 2, which lacks both the node A and first arrow. This diagram would be a natural way to represent a vision of social science – as a sociological atomism – that sociologists often contrast themselves with. An atomist begins with ahistorical assumptions about human nature (as in J. S. Mill's idea of social science) or with purely the stipulated preferences or tastes of individuals (as in sociologists' stereotype of mainstream economics), and proceeds to explain social outcomes. In contrast, sociology is interested in the social factors influencing individuals and their behavior. This does not imply that sociologists would assume that all influences are social, but it is typical of them to assume that social influences (during the life-course of an individual) are of crucial importance. One of Coleman's motivations when introducing the diagram was to use it to show that sociologists have been overly occupied with macro-to-micro influences at the cost of attending to micro-to-macro influences. However, he later commented that "Were I writing the book over again, I would give considerably more attention to this macro-micro relation" (Coleman 1993: 63). In other words, he thinks that both are important.

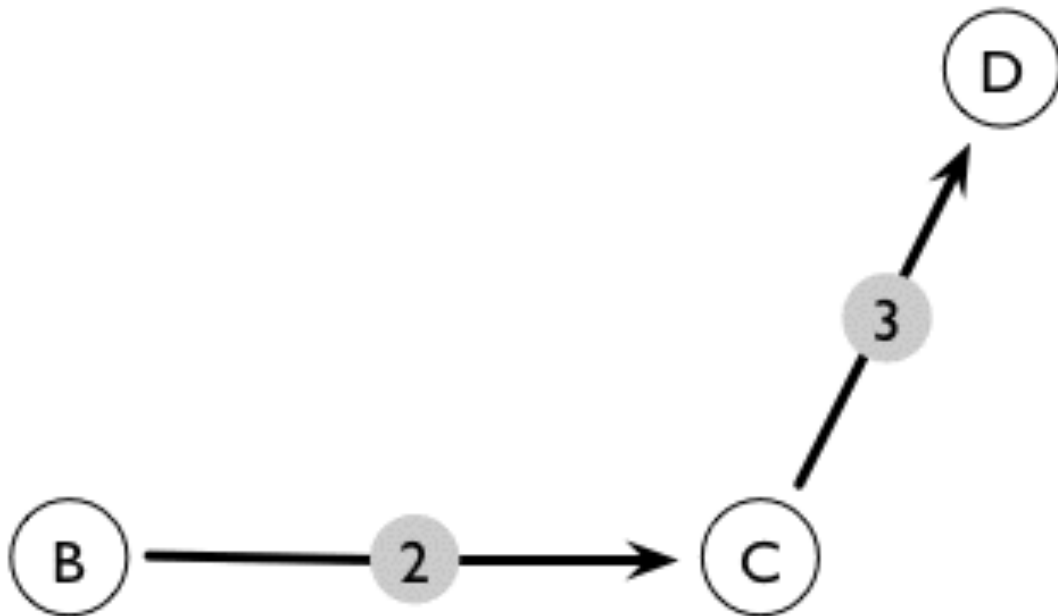


Figure 2. The atomistic version of the diagram

As suggested above, the macro social facts – the A-facts – should be considered broadly. In Coleman’s own work they are quite diverse. For example, there are such things as, for example, “corporate action or social policy” (1990: 646), legislation creating job-training programs (1987: 167, 1990: 643), “improved social conditions” (1987: 156, 1990: 10, 478), “dissatisfaction” (1986, 1321), Protestant religious doctrine (1986: 1322, 1987: 155, 1990: 8), changes in the birth rate (1987: 160), collective fear created by the sound of a fire alarm in a crowded theater (1986: 1321, 1987: 161-162), the “educational level of society” (1987: 165-166), sets of alternatives to collective decisions (1987: 169), the means of production (in Marxist theory) (1986: 23), and the distribution of resources (1992: 121). What is common to these diverse examples is that they all refer to social facts that are “larger” than individuals. While it is impossible to see all of them as belonging to some unique or comprehensive “social level,” their common characteristics are apparent when we think of the micro-macro contrast in terms of scale (Ylikoski 2012, 2014). In this view, macro facts can be of different “sizes”: from local social interactions to whole social systems, thus there is no need to assume the existence of some fixed macro (or meso) levels.

One advantage of thinking macro in terms of scale is that it makes it unnecessary to conceive of macro-to-micro influences in terms of “downward causation.” This metaphorical notion laden with strong metaphysical assumptions can be replaced with the down-to-earth idea of large-scale social facts influencing smaller-scale facts about individuals (or other agents). In terms of difference-making, it is a question of how changes in large-scale facts bring about changes in smaller-scale facts. Furthermore, it is easy to see that such dependencies are contingent on the relevant background conditions (structural, institutional, cultural, etc.): if the background conditions were different, the dependence would be different or absent.

However, there is a complication related to the explanatory dependences between A- and B-facts. While Coleman seems to think that these dependencies are straightforwardly causal, this cannot be the case. Consider Coleman’s favorite example, Max Weber’s Protestant ethic thesis. In Coleman’s reconstruction we have the Protestant religious doctrine (the A-fact) influencing the values of individuals (the B-fact). (Weber did not talk about “values” in the sense of Coleman’s use of the term. However, I will follow Coleman’s terminology to keep the example simple.) In the case of an individual person we could say that replacing the Catholic religious doctrine with the Protestant doctrine causally changes a person’s values. The doctrinal change comes first, or at least it starts first, and the changes in values follow. However, in Coleman’s analysis the adoption of Protestant doctrine is assumed to be a macro fact. The idea is that when the relevant groups adopt Protestantism, their members’ values change. While it is possible that some value changes are actually causal consequences of the conversion to Protestantism, it seems that these changes are not the main focus of Weber’s analysis. At least some of the relevant values have a more intimate relation to the doctrines of Protestant sects. (For Weber, religious doctrines were an important source of evidence concerning these values.) Thus one could say that the group’s adoption of the religious doctrine partly *consists* of its members adopting certain new values. While there is more to

the Protestant doctrine than these values, they are a part of it. Thus it possible to say that as a consequence of the group's adoption of the Protestant doctrine its members' values changed, but one cannot say that the adoption of the doctrine caused the change in values.

It might be possible to fiddle with these concepts in a way that would make this example causal. After all, Coleman's discussion of the case is so sketchy that there is ample room for alternative interpretations. However, the problem is more general. Consider how the demographic change caused by war affects the number of potential marriage partners available to women, how laws allowing same-sex marriage are changing the opportunities of same-sex couples to arrange their legal relationship, or how the improved educational level of society is related to the education of individuals. In all of these cases it is plausible to say that the cited A-facts partially consist of the mentioned B-facts. However, at the same time, some of the relevant consequences might well be causal. Thus it seems that A-B explanatory dependencies are based on (various) mixtures of causal and constitutive relations (Ylikoski 2013). This makes things more complicated, but is this a problem?

Let us look at this in more detail. Consider the following two processes:

- * the macro process A-A* (the adoption of the Protestant doctrine within a community)
- * the micro process B-B* (changes in the values (or life-orientation) of the individual community members)

What are the relations between these two processes? First, we can say that the A-A* process *involves* the B-B* process. However, it involves more, so we cannot simply say that the B-B* process constitutes the A-A* process. Second, we can say that bringing about A-A* would also bring about B-B*. We could also say that the A-A* process would bring about the end-state B*. Third, we cannot say that A-A* is the cause of B-B*, nor can we say that the latter is logically

implied by the former. Finally, it is possible that B* might come about later than A*. This would provide evidence of real causal processes being involved in the chain of changes. Based on these observations we can conclude that the A-B* relation works pretty much like a causal relation, although it involves a mix of both causal and constitutive relations. This complication does not therefore challenge the intelligibility or coherence of the thinking underlying the diagram, a conclusion that is further supported when we realize that the real *explananda* in the diagram are C-facts and D-facts. Whatever the relation is between A- and B-facts, their relation to C-facts is causal.

4.2. Theory of action

The arrow between B and C captures the role of the theory of action in sociology. It bridges between B-states and C-outcomes. As already noted, Coleman has a very liberal notion of agency. He is quite ready to accept non-individual agents, at least for some social purposes (1986: 1312, 1990: 2-5). This is not an isolated feature of his thinking, for one of his major sociological claims is that various forms of corporate agency (corporations, states, unions, parties, etc.) have significantly changed our societies over the last two centuries. Along with many other methodological individualists (Ylikoski 2016), Coleman is more concerned with agency than individuality itself. However, as Coleman does not provide a discussion of non-individual agency in the context of the diagram, I will not discuss it further in this paper.

Another important point is related to Coleman's view of the division of labor between the social sciences and psychology. He says: "a natural stopping point for the social sciences (although not psychology) is the level of the individual - and that, although an explanation which explains the behavior of a social system by the actions and orientations of some entities between the system level and the individual level may be adequate for the purpose at hand, a more fundamental explanation based on the actions and orientations of individuals is more generally

satisfactory” (1990: 4). Here Coleman expresses the idea that social scientific explanations bottom out at the level of individual action. The idea is that social scientists can take it for granted that there are intentional agents, but this is not to be treated as an ultimately unexplainable fact. However, it is the task of psychology (and neurosciences) to explain how we can have these capabilities.

An implication of this division of scientific labor is that social scientists do not have to focus on all of the complex details of individual agency. The task of sociology is to understand complex larger-scale social facts, so it should operate with relatively simple schemes of action explanation.

Coleman identifies an important trade-off: “it is especially important that the individual-action component remain simple. [...] a trade-off between complexity in the other two components and complexity in this component must be made if the overall theory is to remain manageable” (Coleman 1990: 19). His message is that the sociological theory of action can be rather shallow, as the explanation of individual cognition is not the task of the social sciences. What is needed is a scheme that allows bridging A-facts to C-outcomes. I think these ideas are important in understanding Coleman’s adoption of rational choice theory as his preferred theory of action. Rational choice theory provides a very simple version of folk psychology that can be adapted for mathematical modeling. While the rational choice model misses many facts about individual psychology, misconstrues other facts, and is all too simple, it is, in Coleman’s judgment, a useful workhorse for sociological theorizing. Naturally, this assessment can be challenged. There is no reason to assume that sociologists must use the same theory of action in all applications. They can adopt a more pragmatic attitude and give up the idea of foundational sociological theory (Hedström & Ylikoski 2014). This more liberal view does not challenge the utility of the diagram. In this view, the relevant psychological theory is chosen on the basis of what kind of B’s and C’s are the most

relevant in the case in hand. In other words, the pragmatic view presupposes something like the diagram.

Another advantage of this pragmatic approach is that it is also more naturalistic. Coleman is quite attached to the “humanistically congenial image of man” provided by the model of purposive action. He believed that it would provide an interface between sociology, everyday reasoning, economics, and law. In the light of more recent developments in psychology and the cognitive sciences, it is clear that Coleman was too optimistic about the prospects of a folk psychology based scheme for the explanation of social behavior. Thus we may have to give up Coleman’s rather strict division of labor between psychology and the social sciences. For many purposes, explanations that appeal to sub-personal cognitive processes might be better than those based on everyday folk psychology. Thus, for example, mechanisms like the stereotype threat might be of great sociological interest, as they help to connect the properties of social contexts and practices to large-scale social outcomes.

The incorporation of insights from the cognitive sciences, when relevant, does not make the diagram irrelevant. Rather, it highlights its importance, as it helps sociologists to keep an eye on real issues, e.g. the macro facts to be explained and the micro-macro transitions, rather than the foundational theory of action. Thus while Coleman might have been attached to some form of intentional fundamentalism (Ylikoski 2012, Ylikoski & Kuorikoski 2016), this is not an essential part of the diagram. The diagram is based on principles of mechanism-based explanation in the social sciences, and when we replace the foundational theory of action with psychological theories based on cognitive mechanisms it reflects the ideals of mechanism-based explanation even more consistently.

4.3. Micro to macro

For Coleman, the third arrow is the most crucial part of the diagram. The main reason for him starting to utilize the diagram was to highlight the challenges posed by the micro-macro transition. According to him, in both sociology and economics not enough attention is given to the complexities of the micro-macro relation. It is common to assume that the macro order is just an aggregate effect of micro facts about individuals. This is not a problem only for individualistic theories. While holistic theories tend to highlight the contextuality and complexity of everything, in practice their micro-macro assumptions are rather simplistic. It is quite often assumed that macro facts simply reflect the relevant micro facts and vice versa. Schelling's (1978) famous checkerboard model shows how this assumption is wrong in even very simple settings. Although the residents do not favor segregation, they might still end up in a highly segregated neighborhood. Similarly, it cannot be inferred from the segregation of the area that the residents favor this state of affair. In a similar way, Granovetter's (1978) threshold models of collective action demonstrate that it does not make sense to assume that the collective activities reflect the average attitudes of the participating individuals. In fact, the averages can be highly uninformative about macro facts. In both examples there is a discontinuity between micro and macro levels; one could also say that their relation is nonlinear.

To make sense of these relations, Coleman introduces the metaphorical notion of "rules of the game." He suggests that the structural arrangements that determine the micro-macro relations "can be conceived as the rules of the game, rules which transmit consequences of an individual's action to other individuals and rules which derive macro-outcomes from combinations of individuals' actions" (1990: 19). Coleman was not able to provide a general definition of the idea, but does give some examples, including such arrangements as the market and other matching processes. Election rules are another example, and we can use them to highlight the difficulties involved in interpreting his idea.

Let us take as our *explanandum* (the D-fact) the outcome of a parliamentary election, and as the C-fact the individual votes cast by the voters. It is clear that the election outcome (the list of candidates chosen to be the representatives) is not a mere aggregation of valid individual votes. One must utilize the full list of legally set election rules (including the information about voting districts) to determine the election outcome. So here we have a clear illustration of the contrast between mere aggregation and the rules of the game. However, if presented in this way, the example gives the impression that the relation between C- and D-facts is one of logical implication: if you take the valid votes and utilize the rules correctly, the election outcome will follow. This is quite far from the causal interpretation of the micro-macro transition. However, we do not have to interpret the rules of the game in this abstract manner. If we start with the actual votes cast by the voters (in contrast to valid votes) we can see the determination of the election outcome as a causal process. In this interpretation, the rules of the game also involve, in addition to formal election rules, the institutional practices involved in the counting. So if systematic differences exist between districts in vote-counting practices, or if electoral fraud occurs, this would also count as part of the rules of the game. If the rules were different, the outcome of this causal process would be different as well. This is not the only way to conceive of the process; for example, if we start with the voting intentions of the voters (rather than the votes they successfully cast), we get a broader set of rules. In this case, systematic differences in voters' opportunities to vote would also count as part of the rules of the game.

The metaphor of "rules of the game" works relatively well with the election example. However, it is less clear how helpful it is for example in conceiving of the missing elements in Weber's historical scenario about the role of the changing economic behavior of religious sects in the formation of modern capitalism. Another problem with the election example is that it is missing an important element of micro-macro dynamics: the processes that transmit the consequences of an individual's

action to other individuals, in other words, the feedback loops that make agents' choices interdependent. (I will return to this issue later in the paper.) The final problem with the example is that analyzing arrow 3 separately is quite arbitrary. Elections occur in a specific institutional context, and they cannot be understood with only such structural facts as who the legitimate candidates are. While these issues are in principle analyzed earlier when making sense of arrow 1, it is clear that, in practice, one cannot fully separate these two elements of the analysis.

We may conclude that Coleman was not entirely successful in capturing the crucial elements of the micro-macro link to the idea of rules of the game. This leaves an opportunity for others to improve his work. However, it might be the case that it is impossible to give an informative general account of the micro-macro transition. It could be that we must proceed by developing an expanding set of informative examples. This is not only a theoretical challenge; it is also a methodological problem. We still have to learn what the best ways to study micro-macro links are. According to Coleman, "good social history makes the transitions between micro and macro levels successfully" (1990: 21). However, he points out that "it is one thing to be able to trace the development of social organization in a particular instance, as a historian might do, and quite another to develop generalizations about such processes. It is still another to construct models of the macro-to-micro and micro-to-macro processes" (1990: 22). This can be read as a blueprint for studying micro-macro relations, as follows:

1. start with ethnography or case studies
2. develop more general (and abstract) accounts of the key mechanisms in the process
3. build formal (perhaps agent-based) models to test and explore the (generative sufficiency of) suggested mechanisms

4. gather evidence about the particular case to discriminate between the alternative mechanism scenarios

5. The diagram misunderstood

With its increasing popularity, various misunderstandings of the diagram have also proliferated. The diagram is often utilized without familiarity with its original use. This is a potential source of misunderstanding. One such influential misunderstanding is the fusion of the Coleman diagram with the philosophy of mind diagram that is used to discuss the possibility of the mental causation of physical outcomes (See Abell, Felin & Foss 2008; Vromen 2010; Opp 2011; Hedóin 2012). I have elsewhere argued that making an analogy with the mental and the physical is a highly misleading way to conceptualize social micro-macro relations, and that the applicability of the metaphor of levels to micro-macro relations is highly problematic (Ylikoski 2014), so I will not repeat these criticisms here. I will confine myself to showing that the Coleman diagram and the mental causation diagram are quite different. Both the relations and the *relata* are different in a way that makes any analogical relations between the two impossible.

A natural way to begin is to make some observations about the way in which Coleman discusses the micro-macro problem. First, he does not discuss micro and macro in terms of comprehensive levels. As the examples above show, what he says about macro is much better understood in terms of scale rather than layers of reality. For him, the micro-macro contrast is relative, it is not attached to any fixed “levels.” The relevant scale of the macro varies with the relevant sociological research question; similarly, the micro is variable, as he accepts the possibility of corporate and other collective agency. Even more importantly, the *relata* in his diagrams are always individual macro and micro facts. In other words, the micro is not the set of properties on which the macro properties supervene. This alone should make one suspicious of the fusion of these two diagrams. The

diagrams represent two very different relations. The metaphysical diagram aims to represent how an (exhaustive) multitude of agents and their relations constitute a social whole, while the Coleman diagram focuses on the question how a multitude of agents and their relations influence an individual. The first presupposes that it makes sense to conceive of the relation between micro and macro in terms of comprehensive and unique levels (Ylikoski 2012), while the Coleman diagram only presupposes that the larger-scale social context can influence individuals (or more generally, agents). Agency plays a crucial role in Coleman's diagram, but it has no place in the metaphysical diagram.

A second set of considerations is related to uses of these diagrams in arguments. Coleman's diagram is intended to be a tool for explicating and presenting sociological explanations. The metaphysics diagram is basically designed to illustrate the problem of mental causation, or more generally, that of "macro" causation. The purpose of the authors utilizing the latter diagram is to either defend or oppose the possibility of macro causation in a situation where the causal closure of the physical (or micro) is assumed to hold. This is quite different from what Coleman is doing. He never discusses the ideal of causal closure of the physical, and there is no indication that would accept this idea, not to mention the idea about causal closure of the "individual level." Thus, an important presupposition of the macro causation debate is missing. It is true that he is not satisfied with simple macro causal claims in sociology. However, he is not expressing a general doubt about the metaphysical possibility of such causal relations. Rather, he is pointing out how difficult it is to justify such causal claims without understanding the underlying mechanisms, and how deficient such claims are from the point of view of the explanatory aims of sociology. Coleman might be committed to the controversial claim that there are no (satisfactory) causal explanations at the macro level, but he never justifies this claim by arguing that such causal relations are metaphysically impossible. His concerns are sociological, not ontological.

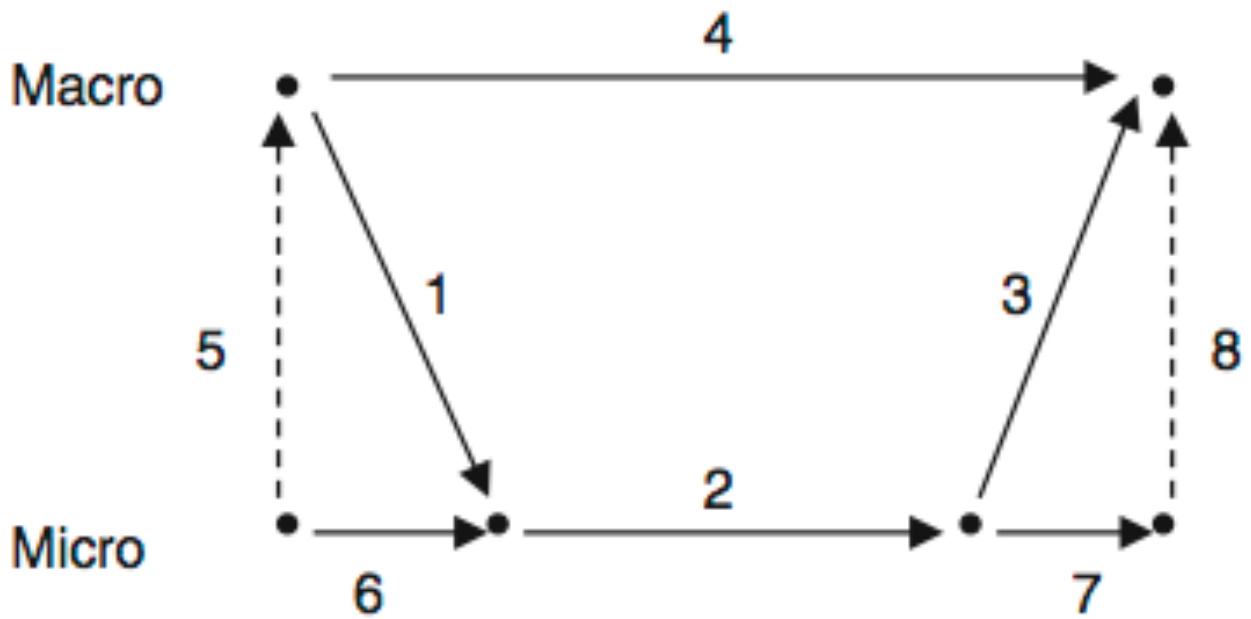


Figure 3. Vromen's (2011: 373) squared diagram

These considerations demonstrate that the two diagrams should not be fused or combined. Jack Vromen (2010) shows with admirable clarity how causal and constitutive questions should be kept separate in the context of mechanistic explanation. However, he ends up combining the two diagrams (Vromen 2010: 373, see the Figure 3) in a way that does not really make sense. First, the metaphysical diagram, but not the Coleman diagram, is based on the controversial presupposition that the idea of comprehensive levels applies to social micro-macro relations. Second, the nodes in the diagram are quite different. In Coleman's diagram they are individual facts, while in the metaphysical diagram they are complete sets of facts. Thus, for example, the Coleman diagram would say that a particular A causes a particular D, while the other diagram would represent the situation where the totality of A's causes the totality of D's. These are very different causal relations. The analogical point holds for all the arrows in the Coleman diagram. The third problem with the combination is that while the Coleman diagram is crucially about agency, the metaphysical diagram has no place for it. For these reasons it is a good idea to keep the diagrams separate.

Another popular misunderstanding of the diagram is related to methodological individualism. Some authors seem to read the Coleman diagram as a summary of his other ideas rather than as an independent tool for sociological thinking. For example, Jepperson and Meyer (2011) and Daniel Little (2012a, b) appear to make this mistake. They begin with Coleman's commitments to methodological individualism and rational choice theory and then use these ideas as cues for interpreting the diagram. Thus, for example, they interpret the downward arrow 1 as an arrow of reduction, and ignore Coleman's rather liberal attitude with respect to corporate agency. This is not the place to discuss how individualistic Coleman's structural individualism actually is (Udén 2001, Ylikoski 2016), or whether it differs in any significant way from Little's methodological localism (Little 2012a); however, the crucial point is that these authors fail to see how Coleman uses the diagram. They read Coleman as belittling the importance of structural and institutional factors when in fact Coleman uses the diagram to *highlight* the importance of these background conditions for social explanation. Similarly, their reductionist reading of the diagram seems to completely miss that the main point of Coleman's discussion is his claim that micro-macro transitions *cannot* be simply aggregative. This is an argument for the importance of "emergence," not its denial. Of course, Coleman also makes it clear that simply labeling something as "emergent" is not very helpful. But this is not a denial of emergence as a phenomenon – rather, a call for theoretical attention to *how* things are emergent (Ylikoski 2014).

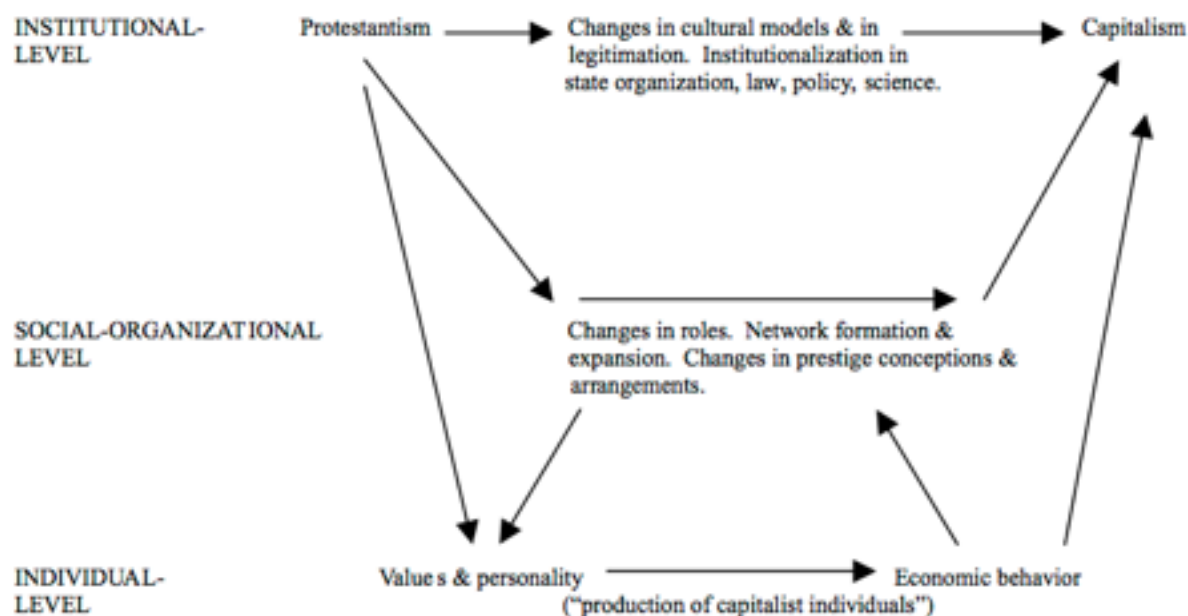


Figure 4. Jepperson & Meyer's (2011: 66) diagram

An additional source of confusion for these authors is the concentration on just one example of Coleman's use of the diagram: Weber's *Protestant Ethic* thesis. It might be that Jepperson and Meyer (2011: 66, see Figure 4) provide a better representation of Weber's theory in their own diagram, but they do not manage to challenge the key theoretical point of the Coleman diagram. The main purpose of Coleman's discussion was not to contribute to Weber scholarship. Coleman used a standard interpretation of Weber's thesis to make a theoretical point. This interpretation might be wrong, but the supposed inadequacy of the theoretical point requires a separate argument. (For an early debate on the adequacy of Coleman's Weber interpretation, see Hernes 1989a,b; Coleman 1989.) The main challenge Coleman puts to social theory is to show *how* various institutional and structural factors work as causes. He is not denying their explanatory value or significance, but is asking how they produce their effects through the agency of individuals (or corporate agents). These explanatory factors are not illegitimate, nor must they be "reduced" to the individual level, but we should understand how they work. So while an analysis at multiple "levels" is often necessary, it is not against the spirit of the diagram. Thus when Jepperson and Meyer posit

various new macro- and meso- “level” factors, they are not making an understanding of the agency foundations of these factors dispensable, but rather multiplying the need for Coleman-type diagrams. The Coleman diagram is quite flexible. It allows for macro facts of various scale, and it might be useful to label some of these “meso.” Thus there is no need for reify these factors to a special “meso-level” that would bring along with it all the conceptual problems of the metaphysics of levels (Ylikoski 2012, 2014).

It is doubtful that Jepperson and Meyer would disagree with the central message here. However, they might have missed the spirit of Coleman’s discussion. His central claim is that micro-macro relations are more complex than sociologists typically assume, and that they require explicit theoretical attention. He is basically saying that sociologists (and other social scientists) are overconfident in their understanding of micro-macro relations. In contrast, Jepperson and Meyer seem to think that microfoundations are a rather trivial philosophical concern that can be easily satisfied. This is precisely the theoretical slumber from which Coleman wanted to awaken sociologists.

6. What is not represented in the diagram?

In this paper I have attempted to explicate how the Coleman diagram works and also defended it at length. However, the diagram also has its limitations. While there is not enough space here for a more extensive discussion of these limitations, the present explication of the diagram can be enhanced by briefly discussing them. Furthermore, understanding these limitations can help us make sense of some of the typical misunderstandings. They could also explain why various modifications of the diagram have not really taken hold. The limitations that I will discuss are all based on things that are not represented in the diagram, but which could be added to it, so they do not pose a fundamental challenge to its usability. However, as we will see, they would make the diagram much more complex and difficult to understand.

The first unrepresented items are the multiple pathways underlying the explanatory dependences. For example, A might influence the relevant B-facts in more than one way. It is also possible that some of these causal pathways have opposing influences. For this reason, arrow 1 basically represents a causal net effect. The multiplicity of causal pathways could in principle be represented by multiple parallel arrows. However, this would, apart from demanding more artistic skill, make the diagram much more complicated without a clear cognitive advantage. The diagram would be more difficult to read and memorize. As the number of needed parallel arrows would in any case vary, it is simpler to use just one arrow. The reader only has to understand the abstraction level of the diagram and recognize that there may be multiple pathways. This might be facilitated by using labels, as introduced by Hedström & Swedberg (1998), for the diagram. Talking about situational, action-formation, and transformational *mechanisms* (in the plural) keeps the reader alert to the possibility of multiple mechanisms.

The second unrepresented but crucial idea is the heterogeneity of agents. The agents represented in the diagram might have different initial beliefs, goals, cognitive schema, routines, and identities. They also hold different social positions, and have different resources and opportunities. There is no reason to assume that the same A-fact would influence them in the exactly the same way. Thus their behaviors (the C-facts) would also be different. The heterogeneity of agents is an important social fact often of crucial explanatory importance. The difficulty is that modeling this heterogeneity is difficult, and social scientists have often tried to solve the problem by employing concepts like representative agent (Hartley 1997). Only recently have agent-based simulation tools made it possible to model how the heterogeneity of agents influences macro outcomes. Nothing in the logic of the diagram implies that we should think that the problem of microfoundations could be resolved by intellectual compromises like representative agents. As in the previous case, this complication could in principle also be represented in the diagram. This could be done having multiple arrows

reaching multiple B-dots, thus implying that there should also be multiple C-dots. However, this would create a very messy diagram. Apart from having huge number of new elements in it, there would be a problem of keeping these arrows separate from the arrows of multiple pathways. It is not clear what the cognitive advantage of these additions would be. The heterogeneity of agents is such a fundamental fact about the social world that it should be remembered without its being included in the diagram.

The causal background conditions of the arrows are the third unrepresented but important issue . While Coleman's point is to draw attention to structural presuppositions, the fact that they are not represented in the diagram makes it easy to forget their importance. Even when one remembers them, one might treat them as fixed background variables. This would lead to an underestimation of their explanatory role, especially in the A-B and C-D transitions. The representation of the causal fields presupposed by the arrow is difficult to incorporate into the diagram. Udéhn's (2001: 305) additional arrows 5 and 6 are difficult to interpret and they seem to break the logic of the diagram. One could also ask whether they are enough to represent all sociologically important causal background conditions. In any case, the domain of application of the Udéhn diagram would be much more limited. It is not a surprise that this modified diagram has not been generally adopted. Apart from these problems representing the causal fields in the diagram, it is also worth asking whether it is necessary to represent them at all. The diagram is not intended to be a complete representation of any sociological explanation, but a tool for developing them. As long as one remembers the rather obvious fact that macro-to-micro and micro-to-macro influences *must* have structural presuppositions, there is no reason to complicate the diagram by adding new elements to it.

Naturally the diagram can also be used as an inspiration for representing the key facts of some specific explanatory narrative (see Erikson 2014). However, in such cases the diagram can, and

should, be modified to fit the particular case. There is no reason to assume that some general enriched version of the diagram could fit all sociological purposes. This highlights the importance of keeping various uses of the diagram clearly distinct.

The fourth and final missing element is the feedback loops that were already mentioned.

Sociologically interesting processes are rarely simple one-shot causal chains that begin with a change in the *explanans* variable and then end with a change in the *explanandum*. In interesting cases, the B-facts are not simply influenced by changes in A-facts but also by the consequences of earlier C-facts. The linear representation of the diagram does not capture this. The idea is the following: the initial A-facts change B-facts about the agents, which in turn influences their behavior (the C-facts). The feedback loop consists of these behaviors in turn influencing the attributes and opportunities of the agents. These changes then influence the agents' behavior and the loop takes another round.

Schelling's (1978) segregation model helps to illustrate this idea. We can interpret the initial setting as a macro change that makes it possible for families to move freely to new neighborhoods. (The randomness of the initial distribution of agents can be interpreted as randomness with respect to the agents' preferences. We could imagine that people were allocated to their houses by some administrative policy.) Following this initial macro change the dissatisfied agents are free to move. Their movements, however, change the composition of both the neighborhood that they are leaving and the neighborhood they are arriving in. Consequently some satisfied agents now face an unsatisfactory situation and maybe some unsatisfied agents become satisfied. This is the idea of the feedback loop. The feedback is based on the fundamental fact of *interdependence* between the agents: what others do affects what others can do and what they want to do. Note that the *explanandum* (D-node) is only the final outcome of this process, so producing it would require multiple rounds of feedback.

Some authors (Abell 2000; Manzo 2007; Abell, Felin & Foss 2008) have attempted to represent the loop by a chain of boats as in Figure 5. However, this does not fully solve the problem, because the loops do not necessarily involve A- and D-facts. Thus chaining boats does not capture the feedback loops that do not go through A and D. This is important because looping can occur on multiple scales. First, there can be a process by which individual agents learn about their environment, opportunities, and other agents. It is possible that they find their new opportunities only after some time, or that they find them appealing after some delay. Second, institutional or structural features that belong to the background causal field can change. They might be included in the generic “macro” of the Manzo diagram, but then there would be a confusion because both generic and specific macro facts would be represented by the same nodes. The third possibility is that macro facts A or D mentioned in the diagram would change as a consequence of agents' actions. Only in this third case does the representation of chained boats work.

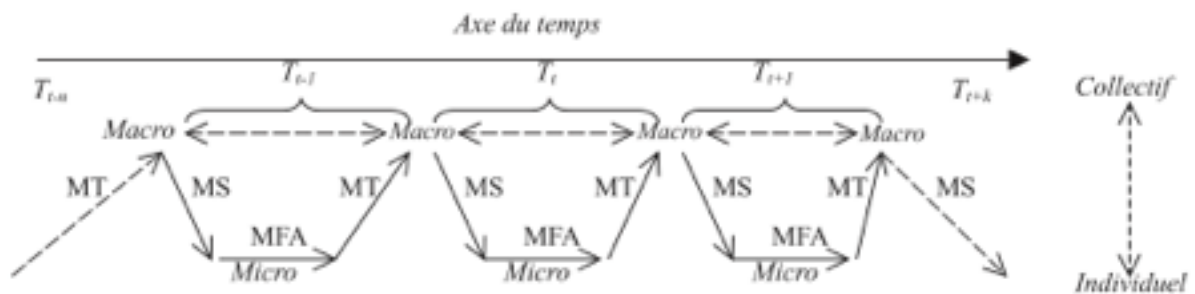


Figure 5. Manzo's (2007: 29) chained boats

An alternative is to represent the feedback as a fifth arrow, as in Figure 6. However, this addition as well requires some interpretive flexibility. The feedback loop can be of various scales, so it can, in principle, go from C to D, or from D to A, or be located at various places on arrows 1 and 3 (reflecting the scale of the feedback process). Representing this variability is difficult without multiple arrows, which would make the diagram very complex and clumsy. However, if the possibility of feedback is only represented by one arrow, the diagram does not become overly

complicated. Given the importance of the issue, this improved version might be a candidate for an updated version of the diagram. Of course, only time will tell whether it will catch on with social scientists.

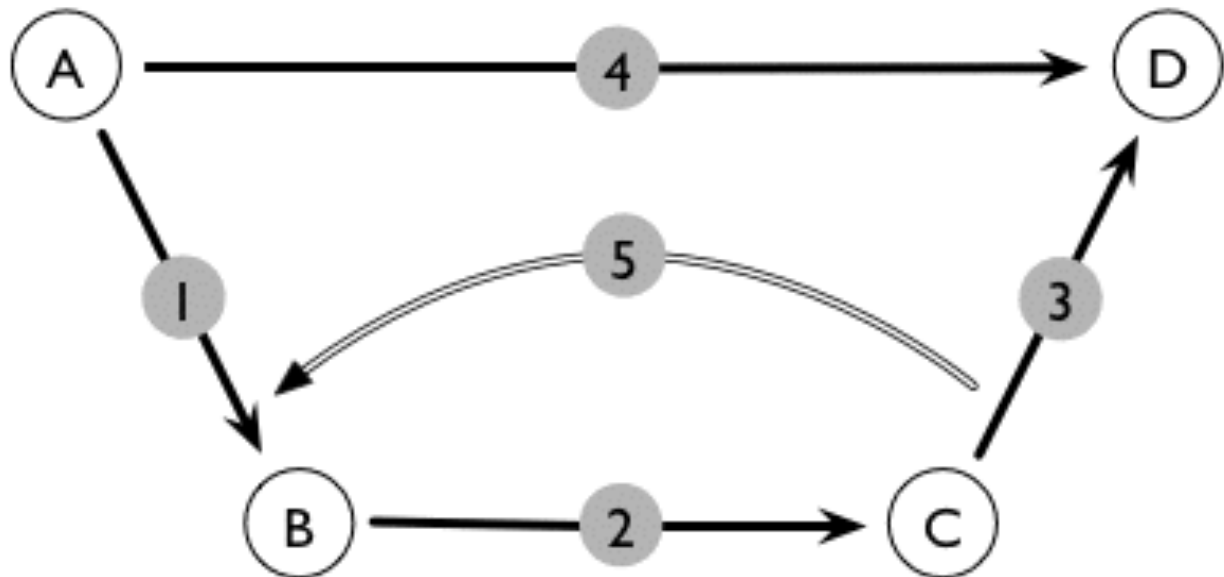


Figure 7. The diagram with feedback

7. Conclusion

This article has provided an extensive commentary on the Coleman diagram. I have argued that, in contrast to many other sociological diagrams, it is not a mere summary of a verbal argument. It is a real tool for thinking about one of the central challenges for sociological theory. Coleman's own discussion of the diagram is very brief, so I have looked at how he uses the diagram and how these uses can be extended to new cases. The discussion of the diagram and its various interpretations show that it is laden with philosophical issues. To discuss the diagram is to discuss some of the central issues in the philosophy of the social sciences. I have defended a causal difference-making interpretation of the diagram that does not commit one to the idea of levels, which is often assumed

in discussions of the micro-macro problem. I think this approach offers the best interpretation of Coleman's own commitments, and also provides the most fruitful basis for further uses of the diagram as a tool for sociological thinking. I have also considered various things that are not included in the diagram and argued that to keep it simple and useful we should not make the diagram too complicated. The only modification that I suggest is the addition of a fifth arrow that would represent a feedback loop from the consequences of action to the structural conditions and individual attributes of agents.

Acknowledgements

This research has received funding from *the European Research Council* under the *European Union's Seventh Framework Programme (FP7/2007-2013)*/ERC grant agreement no. 324233, *Riksbankens Jubileumsfond* (DNR M12-0301:1), and the *Swedish Research Council* (DNR 445-2013-7681 and DNR 340-2013-5460).

References

- Abell, Peter 2000: Sociological theory and rational choice theory. In Turner (ed): *The Blackwell Companion to Social Theory 2nd ed.* Blackwell: 223-244.
- Abell, Peter, Teppo Felin & Nicolas Foss 2008: Building Micro-Foundations for the Routines, Capabilities, and Performance Links. *Managerial and Decision Economics* 29: 489-502.
- Barbera, Filippo 2007: A Star is Born? The Authors, Principles and Objectives of Analytical Sociology. *Papers: Revista De Sociologia* 80: 31-50.
- Beach, Derek and Rasmus Brun Pedersen 2013: *Process-Tracing Methods. Foundations and Guidelines.* Ann Arbor: The University of Michigan Press.

- Coleman, James S. 1986: *Social Theory, Social Research, and a Theory of Action*. *American Journal of Sociology* 91: 1309-1335.
- Coleman, James S. 1987: *Microfoundations and Macrosocial Behavior*, in Alexander, Giesen, Münch & Smelser (eds.): *The Micro-Macro Link*. University of California Press: 153-173.
- Coleman, James S. 1989. *Weber and the Protestant Ethic. A Comment on Hernes*. *Rationality and Society* 1: 291-294.
- Coleman, James S. 1990: *Foundations of Social Theory*. Cambridge, MA: The Belknap Press.
- Coleman, James S. 1992: *The Vision of Foundations of Social Theory*. *Analyse & Kritik* 14: 117-28.
- Coleman, J. S. 1993. "Reply to Blau, Tuomela, Diekmann and Baurmann." *Analyse & Kritik* 15: 62-69.
- Erikson, Emily 2014: *Between Monopoly and Free Trade: The English East India Company, 1600-1757*. Princeton University Press.
- Granovetter, M. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83: 1420-43.
- Hagan, John & Wenona, Rymond-Richmond 2009: *Darfur and the Crime of Genocide*. Cambridge University Press.
- Hartley, James 1997: *The Representative Agent in Macroeconomics*. Routledge.
- Hédoin, Cyril 2012: *Linking institutions to economic performance: the role of macro-structures in micro-explanations*. *Journal of Institutional Economics* 8: 327-349.
- Hedström, Peter & Richard Swedberg 1996: *Rational Choice, Empirical Research, and the Sociological Tradition*. *European Sociological Review* 12: 127-46.

- Hedström, Peter & Richard Swedberg (eds.) 1998: *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.
- Hedström, Peter 2005: *Dissecting the social. On the principles of analytical sociology*. Cambridge: Cambridge University Press.
- Hedström, Peter & Petri Ylikoski 2010: Causal Mechanisms in the Social Sciences, *Annual Review of Sociology* 36: 49-67.
- Hedström, Peter & Petri Ylikoski 2011: Analytical Sociology, in SAGE Handbook of Philosophy of Social Science (edited by I. Jarvie and J. Zamora-Bonilla), 386-398.
- Hedström, Peter & Petri Ylikoski 2014: Analytical Sociology and Rational Choice Theory, in *Analytical Sociology: Norms, Actions and Networks* (edited by Gianluca Manzo), Wiley: 57-70.
- Hernes, Gudmund. 1976: Structural change in social processes. *American Journal of Sociology* 82(3): 513–547.
- Hernes, Gudmund 1989a: The Logic of The Protestant Ethic. *Rationality and Society* 1: 123-162.
- Hernes, Gudmund 1989b: Response to Coleman. *Rationality and Society* 1: 295-300.
- Janssen, Maarten 1993: *Microfoundations: A Critical Inquiry*. Routledge.
- Jepperson, Ronald, & John W. Meyer 2011: Multiple levels of analysis and the limitations of methodological individualisms. *Sociological Theory* 29, 54-73.
- Lindenberg, Siegwart 1981: Erklärung als Modellbau, in W. Schulte (Ed.), *Soziologie in der Gesellschaft*. Bremen, Germany: Zentraldruckerei der Universität: 20-35.
- Little, D. 2012a. Analytical Sociology and the Rest of Sociology. *Sociologica* 1/2012: 1-47.

- Little, Daniel 2012b: Explanatory autonomy and Coleman's boat, *Theoria* 74, 137-151.
- Lynch, M. 1991: Pictures of Nothing? Visual Construals in Social Theory. *Sociological Theory* 9: 1-21.
- Manzo, Gianluca 2007: Progrès et «urgence» de la modélisation en sociologie. du concept de «modèle générateur» et de sa mise en Œuvre, *L'Année sociologique* 57: 13-61.
- McClelland, David 1961: *The Achieving Society*. Princeton: Van Nostrand Co.
- Mills, C. Wright 1959: *The Sociological Imagination*. New York: Oxford University Press.
- Opp, Karl-Dieter 2011: Modeling Micro-Macro Relationships: Problems and Solutions. *The Journal of Mathematical Sociology* 35: 209-234.
- Sampson, Robert 2012: *Great American City. Chicago and the Enduring Neighborhood Effect*. University of Chicago Press.
- Raub, Werner, Vincent Buskens & Marcel A. L. M. van Assen 2011: Micro-Macro Links and Microfoundations in Sociology. *The Journal of Mathematical Sociology* 35: 1-25.
- Raub, Werber & Thomas Voss 2016: Micro-Macro Models in Sociology: Antecedents of Coleman's Diagram, manuscript.
- Rohlfing, Ingo 2012: *Case Studies and Causal Inference*. Basingstoke: Palgrave Macmillan.
- Schelling, Thomas C. 1978: *Micromotives and macrobehavior*. London: W. W. Norton.
- Thornton Patricia H., William Ocasio & Michael Lounsbury 2012: *The Institutional Logics Perspective. A New Approach to Culture, Structure, and Process*. Oxford University Press.
- Turner, Charles 2010: *Investigating Sociological Theory*. SAGE.

Udén, Lars 2001: *Methodological individualism: Background, history, and meaning*. London: Routledge.

Vromen, Jack 2010: MICRO-Foundations in Strategic Management: Squaring Coleman's Diagram, *Erkenntnis* 73: 365-83.

Ylikoski, Petri 2011: 'Social Mechanisms and Explanatory Relevance', in *Analytical Sociology and Social Mechanisms* (edited by P. Demeulenaere), Cambridge University Press, 154-172.

Ylikoski, Petri 2012: Micro, Macro, and Mechanisms, *The Oxford Handbook of Philosophy of the Social Sciences* (edited by H. Kincaid), Oxford University Press, 21-45.

Ylikoski, Petri 2013: Causal and constitutive explanation compared, *Erkenntnis* 78, 277-297.

Ylikoski, Petri 2014: Rethinking Micro-Macro Relations, in *Rethinking the Individualism-Holism Debate. Essays in Philosophy of Social Science* (edited by Finn Collin & Julie Zahle), Springer, 117-135.

Ylikoski, Petri 2016: Methodological Individualism, in *Routledge Companion to Philosophy of Social Science* (edited by Lee McIntyre and Alexander Rosenberg), Routledge.

Ylikoski, Petri & Emrah Aydinonat 2014: 'Understanding with Theoretical Models', *Journal of Economic Methodology* 21: 19-36.

Ylikoski, Petri & Jaakko Kuorikoski 2016: Self-interest, norms, and explanation, in *Normativity and Naturalism in the Social Sciences* (edited by Mark Risjord), London: Routledge: 212-229.