

Understanding the Coleman boat

Petri Ylikoski

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 chapter, I will 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, the diagram has become a popular representation of the micro-macro challenge in analytical sociology, and also outside of it.

The diagram is interesting 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 relationship to smaller-scale social events and processes (like individual behavior, social interactions etc.) provides a thorny set of problems. The diagram offers a way to think about these micro-macro relationships.

Second, the diagram is a scarce visual representation in sociological theory. Most diagrams by sociological theorists are mere summaries of verbal argument or otherwise primitive compared to the visual representations developed in other sciences (Lynch 1991; Turner 2010; Swedberg 2016). In contrast, the Coleman diagram is a representation that can serve as a cognitive tool for thinking, a diagram for theorizing (Swedberg 2016). It is “a genuine tool of production,” as Wright Mills (1959, p. 213) puts it. The diagram provides a systematic scheme for articulating social explanations and their presuppositions. The contents of the scheme could, in principle, be verbally articulated, but the easy to remember 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 claim to have a theoretical understanding of the social phenomenon.

However, an abstract diagram can be interpreted in many different ways. Even a cursory look at various interpretations, modifications and applications of the diagrams demonstrate that there is no general agreement about what the diagram represents. For example, many critics have taken the diagram as a summary of Coleman's theoretical ideas. Thus, it has been interpreted as a statement of 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 mistaken. They reduce the diagram to a (bad) theory picture (Swedberg 2016) that loses the productive contribution the diagram can make. In the following I will discuss the diagram as a cognitive tool that is independent of Coleman's other theoretical commitments. Also, people who might disagree with Coleman on methodological individualism or rational choice theory can make productive use of the diagram. It is a tool for thinking, not a metaphysical or theoretical statement.

This chapter aims 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 used 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 chapter is the following. Section 2 will describe the diagram and its elements and how Coleman used the diagram in his work. Section 3 follows the arrows of the diagram and discusses their interpretation and presuppositions. It also introduces an additional arrow to the diagram to make the micro-macro transition more understandable.

2. THE NODES AND ARROWS

This chapter will focus on how Coleman used the diagram and how it can be improved. Later authors have given the diagram many alternative interpretations that I cannot discuss here.¹ To facilitate the discussion, I use capital letters to refer to nodes of the diagram and numbers to the arrows, as shown in Figure 1.1. Coleman himself did not use letters to name the nodes, as in his diagrams the nodes always contained a reference to relevant causal factor or variable. Similarly, the numbering of the arrows is not present in Coleman's original diagrams, but their use has become standard practice. However, the numbering varies between authors. In some versions of the diagram the arrow between nodes A and D are missing. This absence does not change the meaning of the diagram. The reason why I have labeled it with a question mark rather than number will become apparent soon.

¹ Coleman did not invent the diagram out of nothing and, regrettably, he did not refer to any of his precursors. One clear precursor can be found in David McClelland's book *The Achieving Society* (1961, p. 47). McClelland's diagram has the same recognizable arrows, and he also uses the same example, Weber's thesis in Protestant Ethic. It is less clear whether McClelland used the diagram in the same way as Coleman. (It is also noticeable that there are conceptual precursors, like Hernes 1976 and Boudon 1981, that used alternative visual means to represent the same ideas.) However, it seems that all the later published uses of the diagram refer to Coleman (or to authors who refer to Coleman). 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 and Swedberg 1996; Hedström and Swedberg 1998; Abell 2000; Udéhn 2001; Manzo 2007; Abell, Felin and Foss 2008; Sampson 2012; Thornton, Ocasio and Loundsbury 2012) However, apart from Hedström's and Swedberg's 1998 mechanistic interpretation, these modifications have not been adopted for more general use. For the history of the diagram see, Barbera 2007 and Raub and Voss 2017.

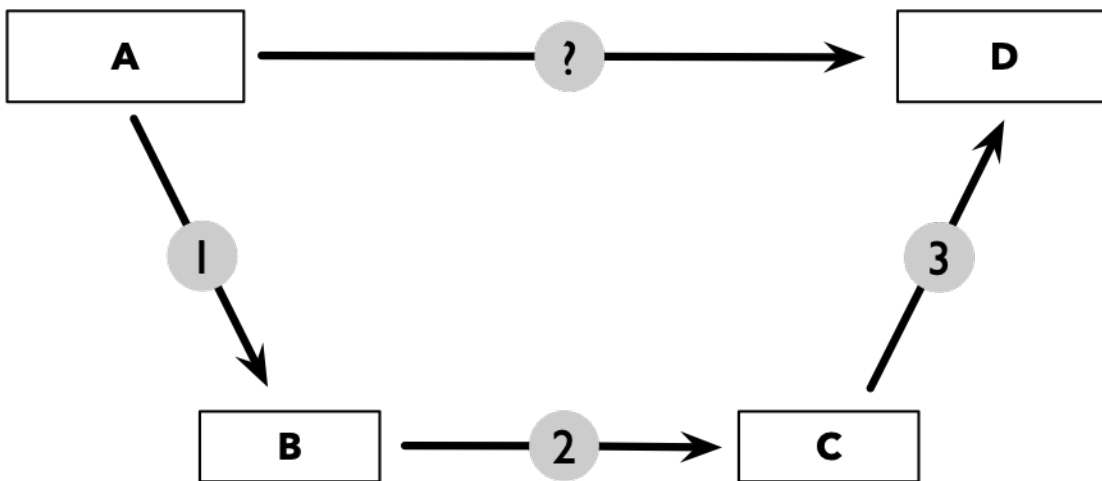


Figure 1.1 The Coleman diagram

2.1 The Nodes

To use the diagram, one has to understand what the nodes and arrows mean. The nodes A and D refer to various sorts of macro social facts that figure as *explanantia* and *explananda* in sociological explanations. A-facts are extra-individual social factors that might be cited as causes of social phenomena and might influence individuals. D-facts are macro-social facts to be explained. These facts can be quite diverse. Among Coleman’s sociological *explananda* are revolutions, economic structures, inequalities, panics, election outcomes and disintegration of families. Coleman’s notion of the macro is flexible: its scale can vary from a two-person social interaction to organizations and nation states (Coleman 1986, 1987, 1990; Manzo 2020). Also, the A-facts can be quite different in their scale. Coleman says that A-facts could be “. . . characteristics of their social environments, ranging from family to friends to larger social contexts” (1990, p. 1). The A-facts are not privileged: in principle, all A’s can (and should) serve as D’s for other explanations. In other words, there are no unexplainable social facts. It is crucial to understand that A and D do not refer to the whole of “macro-level” – assuming that it is even coherent notion – but only to specific facts relevant for the hypothesis under consideration.

The node B refers to properties of agents and their situations. Most of Coleman’s examples of B-facts are about values, opportunities or interests, but this set can also be extended to analogical theoretical concepts like beliefs, desires, goals, values, preferences, motives, emotions, habits (or habitus), routines, scripts, heuristics, cognitive schema or identities (social and psychological), as all these notions play a similar role in sociological explanation. They are theoretical concepts used in the explanation of human behavior, and their purpose is to mediate between social influences and individual behavior. Some of these notions might ultimately be too vague to be useful in social

science, but the evaluation of these concepts is not a concern of this chapter. The B-facts are important because they explain behavior (C-facts), and they mediate the influence of A-facts on behavior. While there might be theoretical reasons to avoid some of these notions, the explanation of action requires that some such theoretical concepts are employed. Sociologists are not primarily interested in explaining B-facts themselves; instead, B-facts are a way to link A-facts to C-facts.

The 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 in explaining macro outcomes (D-facts). Coleman allows that B and C facts are not about individual persons. They could “. . . be institutions within the system or subgroups” (1990, p. 2). They could also be corporate agents. Coleman’s notion of micro is flexible, as is his notion of the macro. This flexibility implies that the diagram cannot be understood as an argument for methodological individualism (contra Jepperson and Meyer 2011). For Coleman, the agency is of primary importance, not individuality (or personhood) of agents (Ylikoski 2017). Thus, he writes, “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, p. 1312).

2.2 The Arrows

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

All of these are straightforward causal claims. “Generation” and “bringing about” are everyday causal notions referring to causal processes. It is also evident that in the second sentence the values are assumed to influence the economic activities causally.

For 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.² My simple suggestion is that the arrows are interpreted in terms of difference-making. Thus, for example, in the case of Arrow 1 some change in A-variable brings about a change in B-variable. This is a kind of minimal explanatory dependency that is easy to conceive in terms of (imagined) causal interventions: changing A to A* would bring about change in B to B*. For example, parliament changes the legislation about parental leaves (the intervention on A-variable), which brings about changes in workers’ opportunities to take care of their small children (the change in 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 these changes were brought about. However, one might not have that information available, or there are multiple causal pathways (mechanisms) at work that might in opposition to each other, so one would still like to know whether there is some sort of net effect on B. In addition, the difference-making interpretation provides an excellent way to conceptualize structural presuppositions of macro-micro and micro-macro claims: they are background conditions – or a causal field (Mackie 1974) – that make the foreground causal claims possible.

It is also useful to think about the arrows in terms of questions they answer. Arrow 1 is associated with the question, How does the macro change affect the agents? Arrow 2 with the question, How does the behavior of agents change? and Arrow 3 with the question: How do changes in behavior bring about macro consequences? The questions can be formulated in various ways, but formulating them in terms of how-questions highlights the interest in the mechanisms of change: it is not enough to know that certain changes are causally related, we also wish to know how the changes happen.

² Coleman himself did not explicitly advocate any specific account of explanation. The Dutch interpretation that relies on covering-law account of explanation and reconstructs the diagram in terms of deductive arguments (see Lindenberg 1981; Raub, Buskens and van Assen 2011 for details). The alternative Swedish interpretation (Hedström and Swedberg 1998), interprets the arrows in terms of causal mechanisms.

Finally, there is the arrow at the top that I have labeled with a question mark. This arrow is sometimes absent from Coleman's own diagrams.³ It is the starting point of the analysis that employs the diagram. It can be an observed or hypothetical connection between two macro variables. In such cases the purpose of the diagram is to help to determine whether the relationship between A and D is genuinely causal and, if it is, why does it hold. However, D might also be unknown; in such cases, the goal is to reason through the diagram and figure out *what could be* the macro consequences of a change in A.

2.3 The Motivations

The motivation to use the diagram is to make macro causal claims more secure by finding the underlying causal mechanisms. For example if the university leadership believes that changes in professors' salaries increase 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. In addition to *the justification of causal claims* the other added value provided by the diagram is *theoretical understanding*. Rather than just citing a relationship between two macro variables, the ideal sociological explanation provides an account 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. In other words it provides an action-based explanation. The diagram shows how such explanations have three analytically distinct explanatory subtasks. For this reason, the diagram is based on the rationale of mechanism-based explanation. 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).

While it cannot be expected that a single study or publication could provide the complete story, the diagram tells us what parts a full sociological explanation would have to have. It provides a skeleton for an ideal explanatory narrative. For this reason it can be utilized as a tool for thinking about particular sociological explanations. Coleman offers many examples of this. First, he uses the

³ Coleman had the arrow in the diagram in his 1987 and 1990 publications, but not in his 1986 and 1992. There is no reason to suppose that this reflects any changes in his thinking.

diagram to *illustrate* specific explanatory ideas – for example the marriage and other matching markets (1987, pp. 159–160) and problems with collective decision-making (1987, pp. 169–71; 1990, pp. 400–2) – and to *compare* competing explanations, for example the explanations for panic in a burning theatre (1987, pp. 161–3, see also 1990, pp. 203–9). More importantly, he employs the diagram to *point out missing crucial elements* in the proposed explanatory narratives. This is the main point of his discussion of Weber’s thesis in *Protestant Ethic* (1987, pp. 154–5; 1990, pp. 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, p. 9)

Finally, Coleman uses the diagram to *articulate structural assumptions of explanations*. Good examples of this are his discussions of the influence of the educational inequality on income inequality (1987, pp. 163–6) and the impact of training programs on youth employment (1987, pp. 166–8; 1990, pp. 642–5). In both cases the diagram is used to trigger articulations of presuppositions of suggested explanations so that their credibility can be evaluated.

The diagram can also be used in thinking about foundational issues in social theory. First, it is useful in highlighting the non-trivial nature of the problem of micro-foundations. Both explanations of C-facts and D-facts involve structural assumptions that are crucial for understanding the explanatory dependencies. Here is Coleman’s key message: the micro-macro relationships are aggregative only in exceptional cases. Second, the diagram can be used to illustrate the role of the “theory of action” in a sociological explanation. While pure BC-explanations fall mainly 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 (direct AD-explanations) are not theoretically satisfactory, and they need to be supplemented by an account of the micro process underlying it. The point of mechanism-based explanation is to show how the causal (or non-causal) relationship between A and D is brought about. This helps us to both justify or criticize the suggested causal connection and have an explanatory understanding of it.

3. THINKING THROUGH THE DIAGRAM

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

3.1 Macro to Micro

Causal influences from macro to micro are a central concern in sociology. One could say that they are a defining feature of a distinctly sociological approach to the social world. This is apparent when we consider an alternative diagram shown in Figure 1.2. This diagram would be a natural way to represent a vision of social science – asociological 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 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 social factors influencing individuals and their behavior. This does not imply that sociologists would assume that all influences are social, but it is typical for them to assume that social influences (during the life-course of an individual) are of crucial importance. One of Coleman’s motivations for introducing the diagram was to show that sociologists have been too occupied with macro-to-micro influences at the cost of attending to micro-to-macro influences.⁴

⁴ However, Coleman later commented, “Were I writing the book over again, I would give considerably more attention to this macro-micro relation” (Coleman 1993, p. 63). In other words, he thought that both are important.

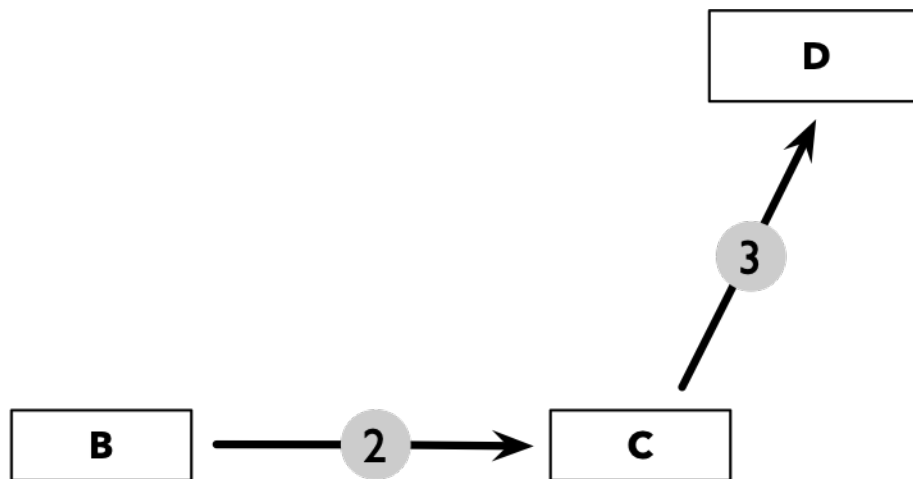


Figure 1.2 Atomistic version of the diagram

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 corporate action or social policy (1990, p. 646), legislation creating a job-training program (1987, p. 167; 1990, p. 643), improved social conditions (1987, p. 156; 1990, pp. 10, 478), dissatisfaction (1986, p. 1321), protestant religious doctrine (1986, p. 1322; 1987, p. 155; 1990, p. 8), change in birth rate (1987, p. 160), a collective fear created by a sound of fire alarm in a crowded theatre (1986, p. 1321; 1987, pp. 161–2), educational level of society (1987, pp.165–6), the set of alternatives for collective decisions (1987, p.169), the means of production (in the Marxist theory) (1986, p. 23), and distribution of resources (1992, p. 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 shared 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 that there is some fixed macro (or meso) level.

An advantage of thinking macro in terms of scale is that it makes it unnecessary to conceive 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 dependencies like this are contingent on relevant (structural,

institutional, cultural etc.) background conditions: if the background conditions were different, the dependence would be different or absent.

However, there is a complication related to explanatory dependencies 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).⁵ In the case of a person we could say that replacing the Catholic religious doctrine with the Protestant doctrine causally changes a person's values. The doctrinal change is first, or at least it starts first, and the changes in values follow. However, in Coleman's analysis, the adoption of the Protestant doctrine is assumed to be a macro fact. The idea is that when the relevant groups adopted Protestantism, their members' values changed. While there may be some value changes that are really causal consequences of the conversion to Protestantism, it seems that these are not changes that are the main focus of Weber's analysis. At least some of the relevant values have a more intimate relationship to doctrines of the Protestant sects. (The religious doctrines were an important source of evidence about these values for Weber.) Thus, one could say that the group's adoption of the religious doctrine partly *consists* of its members adopting specific new values (Ylikoski 2013). While there is more to the Protestant doctrine than these values, they are part of it. Thus, one can 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 turn this example causal. After all, Coleman's discussion of the case is so sketchy that there is a lot of 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 for women, how the law allowing same-sex marriage changes the opportunities of same-sex couples to arrange their legal relationship, or how improved the educational level of society is related to the education of individuals. In all 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 important consequences might well be causal. Thus, it seems that A-B explanatory dependencies are based on (various) mixtures of causal and constitutive relationships (Ylikoski 2013).

⁵ Weber did not talk about 'values' in the sense Coleman uses the term. However I will follow Coleman's terminology to keep the example simple.

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

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

What are the relationships 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 even say that the process A-A* 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, B* may be later than A*. This would provide evidence there are real causal processes involved in the chain of changes. Based on these observations we can conclude that the A-B* relationship works pretty much like a causal relationship, although it consists of a mix of causal and constitutive relationships. Thus, this complication does not challenge the intelligibility or coherence of the thinking underlying the diagram. This conclusion is further supported when we notice that the real *explananda* in the diagram are C-facts and D-facts. Whatever is the relationship between A- and B-facts, the link of A-facts to C-facts is causal.

Another complication are the things that are not represented in the diagram. All the arrows in the diagram abstract away from causal details: they have background conditions, and they might consist of multiple causal pathways. Let us first discuss the possibility of multiple pathways. For example, A-facts 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 net causal effect. The multiplicity of causal pathways could, in principle, be represented by multiple parallel arrows. However, this would, apart from demanding more from artistic skill, make the diagram much more complicated without clear cognitive advantage. The diagram would be more challenging to read and memorize. The reader just has to understand the abstraction level of the diagram and recognize that there might be multiple pathways.

Presuppositions of the dependencies represented by the arrow are the second thing that the diagram does not show. As suggested above the dependencies hold against the background of the causal field. What this causal field consists of is often a significant challenge for a theorist, and it poses problems for theoretical generalization as many of the presuppositions are structural. As with

the case of multiple pathways the reader has to recognize the role of causal background conditions without them being represented. 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 multiple new elements.

3.2 Theory of Action

The arrow between B and C captures the role of theory of action in sociology. It bridges between B-states and C-outcomes. As already noted Coleman has quite a liberal notion of agency. He is quite ready to accept non-individual agents, at least for some explanatory purposes (1986, p. 1312; 1990, pp. 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 2017), Coleman is more concerned with the agency than with 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 chapter.

Another important point is related to the Coleman's view of the division of labor between 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, p. 4)

Coleman expresses here 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. However, 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.

This division of scientific labor implies that social scientists do not have to focus on all intricate 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, p. 19)

His message is that the sociological theory of action can be relatively shallow from the point of view of psychology, 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. This perspective helps to understand why Coleman adopted the rational choice theory as his preferred theory of action. The 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, gets other facts wrong, and is all too simple, it is, in Coleman's judgment, a useful workhorse for sociological theorizing. This assessment can be challenged. There is no reason to assume that sociologists have to use the same theory of action in all contexts. They can adopt a more pragmatic attitude and give up the idea of foundational sociological theory of action (Hedström and Ylikoski 2014).

Another advantage of the more pragmatic approach is that it is also more naturalistic. Coleman (1986) is 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 light of more recent developments in psychology and cognitive sciences, it is clear that Coleman was too optimistic about the prospects of a folk psychology-based scheme. We may have to give up Coleman's rather strict division of labor between psychology and social sciences. For many purposes, explanations that appeal to sub-personal cognitive processes might be better than those based on everyday folk psychology. The incorporation of insights from cognitive sciences, when relevant, does not make the diagram irrelevant. Instead, it highlights its importance, as it helps sociologists to keep their 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), 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.

The third element that the diagram does not represent consists of 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 precisely the same way. Thus, their behaviors (the C-facts) would also be different. The heterogeneity of agents is an important social fact, and it is often of crucial explanatory importance. Modeling of this heterogeneity is intricate and social scientists have often tried to solve the problem by employing concepts like the representative agent (Hartley 1997). Only recently have agent-based simulation tools made it possible to model how the heterogeneity of agents influences the macro outcomes (Manzo 2020). Nothing in the logic of the diagram implies that we should think that intellectual compromises like representative agents could solve the problem of micro-foundations. Like background conditions and multiple pathways the heterogeneity of agents is a fundamental fact about the social world that should be remembered when using the diagram.

3.3 Micro to Macro

For Coleman, the third arrow is the most crucial part of the diagram. The main reason he started to utilize the diagram was to highlight the challenges posed by the micro-macro transition. According to him, both in sociology and economics, not enough attention is given to the complexities of the micro-macro relationship. It is common to assume that macro order is just an aggregate effect of micro facts about individuals. Notice that this is not only a problem for individualistic theories. While holistic theories tend to highlight contextuality and complexity, 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 favour segregation, they might still end up in a highly segregated neighbourhood. Similarly, it cannot be inferred from the segregation of the area that the residents favour this state of affairs. Similarly, 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 participating individuals. In fact, the averages can be highly uninformative about macro facts. Even these elementary examples show that there is a discontinuity between micro and macro.

To make sense of these relationships, Coleman introduced the metaphorical notion of "rules of the game." He suggests that the structural arrangements that determine the micro-macro relationships:

. . . 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, p. 19)

Coleman did not provide a general definition of the idea, but gave some examples such arrangements as the market and other matching processes. However, he did not spell out the "rules" of these examples. Furthermore, the metaphor is not very helpful in conceiving the missing elements in Weber's historical scenario. The crucial aspects of the micro-macro link are not captured with the idea of rules of the game. This leaves an opportunity for others to improve his work.

The arrows of the diagram give it a dynamical nature, as the observer follows the arrows, s/he also follows the explanatory narrative consisting of the chain of arrows. However, this representation has the disadvantage of being linear and giving an impression of the one-shot process. 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 most interesting cases the B-facts are not merely influenced by changes in A-facts, but also by the consequences of earlier C-facts.

This is the fourth missing element that is not captured by Coleman's original diagram. 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, in turn, influence the agents' behavior, and the loop takes another round.⁶ This feedback can be represented by adding an extra arrow to the diagram.⁷ Now the unfolding social process can be analyzed as a looping process, and D is the outcome of this process.

The endpoint of the Arrow 3 is purposefully ambiguous. The behavioral consequences could either affect other agents directly or via the institutional and structural conditions. This is an empirical question; the purpose of the arrow is to remind the researcher that it is important to look for these

⁶ The idea of a loop is not entirely new. Some authors (Abell 2000; Manzo 2007; Abell, Felin and Foss 2008) have attempted to represent the loop by a chain of boats (where each D is followed by A of the next boat). However, this does not fully solve the problem because the loops do not necessarily involve A and D facts. Thus, chaining boats do not capture those feedback-loops that do not go through A and D.

⁷ To keep the order of the arrows logical, I have changed the number of the micro-to-macro arrow.

feedback loops as they are often crucial to the process of social change. The looping can occur at multiple scales. First, there can be a process by which individual agents learn about their environment, opportunities and other agents. They may find their new opportunities only after some time, or they find them appealing after some delay. Second, institutional or structural features that belong to the background causal field can change.

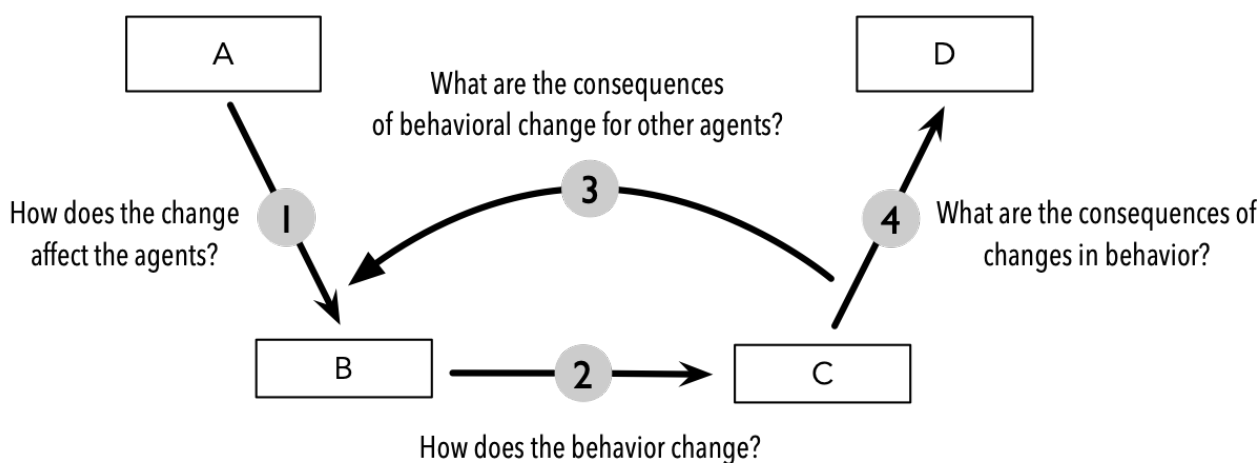


Figure 1.3 Coleman diagram with a feedback loop

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 agents' preferences. We could imagine that people were distributed to their houses by some administrative policy.) After 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 where they are arriving. This has the consequence that some satisfied agents are now facing an unsatisfactory situation, and maybe some unsatisfied have become satisfied. This is the idea of the feedback loop. The feedback is based on the fundamental fact of *interdependence* between the agents: what an agent does affects what others can do and what they want to do.⁸

⁸ It might be useful to distinguish between two forms of interdependence. *Strategic interdependence* is familiar from the game theory. It deals with the interdependence of choices: the consequences (and rationality) of choice are dependent on what others choose. The feedback loop describes *causal interdependence*: the agent's behaviors influence other agents' later beliefs, desires or opportunities (or other B-facts).

Notice that the *explanandum* (D-node) is only the final outcome of this process, so producing it could require multiple rounds of feedback.

However, the Schelling example does not fully capture the potential of the diagram. In many interesting cases the result of the process is a *structural change*, e.g. the rules of the game change (Hernes 1976). Consider the effects of US News law school rankings studied by Espeland and Sauder (2016). The introduction of the ranking can be regarded as the initial macro change, as it was a large-scale change affecting all relevant agents (Ylikoski 2012). According to Espeland and Sauder, the ranking provided a cognitive map for applicants who were interested in comparing law schools (Arrow 1). With the help of it the earlier very complicated choice was simplified dramatically: it was not possible to compare directly all schools to each other without presupposing any expertise about legal education. Thus, the wide availability and the easy use of the rankings made them a very influential factor in choosing a law school (Arrow 2).

To capture the unintended macro consequences of the rankings, we have to focus on the *interdependence of the agents*; in other words, the impact the changes in the applicants' behavior had on other agents. This interdependence is captured by the third arrow representing the causal loops by which the agents influence each other through their behavior. Thus, while most deans initially regarded the *US News* rankings as both incompetent and harmful, their adoption by the applicants made the rankings relevant to them. When the deans started to worry about the rankings, both the content and evaluation criteria of the work by people working in admission and career changed. Similarly, once applicants began to make decisions based on rankings, the employers could start using the rankings as a source of information about the quality of job applicants. Furthermore, when the employers began to use the rankings (or were widely believed to do so), the school rankings became an increasingly important part of the market value of the student's degree. This, in turn, increased the pressure felt by the deans and created a dynamic in which the importance of the rankings was amplified as an increasing number of groups started to react to them.

The created dynamic is crucial for the long-standing changes in legal education produced by the rankings (Arrow 4). The attempts to improve the school's standing involved, among other things, changes in how students were admitted, how the schools utilized their resources, to whom the schools targeted their marketing, and how the career services advised fresh graduates. These changes, in turn, generated structural changes in US legal education; for example, the transformation of the legal curriculum changed the composition of the student body, and the

increasing similarity and status competition between the schools (Espeland and Sauder 2016, Chapter 7).

With the added arrow the diagram gets more dynamical as it no longer gives an impression of a one-shot process. With the feedback arrow in place it is much harder to assume that the micro-macro transition is just a simple process of aggregation. It also highlights how thinking with the diagram is a natural way of approaching the complexity of social processes: heterogeneity and interdependence of agents, sensitivity to the background conditions, and various sorts of feedback processes. While only the last one is “officially” represented in the diagram, the questions about others are its natural corollaries.

4. CONCLUSION

This chapter has been extensive commentary on the Coleman diagram. I have argued that, in contrast to many other sociological diagrams, it is not a mere summary of verbal argument. It is a tool for thinking about one of the central challenges for sociological theory. Coleman’s discussion of the diagram was very brief, so I have looked at how he uses the diagram and how these uses can be extended to new cases. The interpretation of the diagram is laden with deep questions about the nature of social theory. As any diagram, it can be given multiple interpretations, and any user of the diagram can use it anyway they find it illuminating.

In my rational reconstruction I have tried to preserve the central idea that the diagram is a tool for thinking. I believe that this was Coleman’s original intention, which is most visible in his 1987 chapter. My reconstruction also highlights four important issues for social theorizing: sensitivity to the background conditions, heterogeneity, interdependence of agents and various sorts of feedback loops. To capture the last, I introduce one significant change to the diagram – the feedback arrow – because the addition improves the usability of the diagram and does not involve a radical departure from Coleman’s thinking.

References

Abell, Peter, (2000), “Sociological theory and rational choice theory”, in Turner (ed.) *The Blackwell Companion to Social Theory*, 2nd ed. Oxford: Blackwell: 223–44.

- Abell, Peter, Teppo Felin and Nicolas Foss (2008), “Building micro-foundations for the routines, capabilities, and performance Links”, *Managerial and Decision Economics*, **29** (6), 489–502.
- Barbera, Filippo (2007), “A star is born? The authors, principles and objectives of analytical sociology”, *Papers: Revista De Sociologia*, **80**, 31–50.
- Boudon, Raymond (1981), *The Logic of Social Action. An Introduction to Sociological Analysis*, London: Routledge and Kegan Paul.
- 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 and Smelser (eds), *The Micro-Macro Link*, Berkeley: University of California Press: 153–73.
- 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 and Kritik*, **14**, 117–28.
- Coleman, James S. (1993), “Reply to Blau, Tuomela, Diekman and Baumann”, *Analyse and Kritik*, **15**, 62–9.
- Espeland, W.S. and Sauder, M. (2016), *Engines of Anxiety Academic Rankings, Reputation, and Accountability*. New York: Russell Sage Foundation.
- Granovetter, M. (1978), “Threshold models of collective behavior”, *American Journal of Sociology*, **83**, 1420–43.
- Hartley, James (1997), *The Representative Agent in Macroeconomics*, London: Routledge.
- Hedström, Peter and Richard Swedberg (1996), “Rational choice, empirical research, and the sociological tradition”, *European Sociological Review* **12**, 127–46.
- Hedström, Peter and Richard Swedberg (eds) (1998), *Social Mechanisms: An Analytical Approach to Social Theory*, Cambridge: Cambridge University Press.
- Hedström, Peter and Petri Ylikoski (2014), “Analytical Sociology and Rational Choice Theory”, in Gianluca Manzo (ed.), *Analytical Sociology: Norms, Actions and Networks*, Chichester: Wiley: 57–70.

- Hernes, Gudmund (1976), “Structural change in social processes”, *American Journal of Sociology* **82** (3), 513–47.
- Jepperson, Ronald and 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.
- Lynch, M. (1991), “Pictures of nothing? Visual construals in social theory”, *Sociological Theory*, **9**, 1–21.
- Mackie, J. L. (1974), *The Cement of the Universe*, Clarendon Press. Oxford.
- Manzo, Gianluca (2007), “Progrès et ‘urgence’ de la modélisation en sociologie. du concept de ‘modèle générateur’ et de sa mise en oeuvre”, *L’Année sociologique* **57**, 13–61.
- Manzo, Gianluca (2020), “Agent-based models and methodological individualism: are they fundamentally linked?” *L’Année sociologique*, **70**, 197–229.
- McClelland, David (1961), *The Achieving Society*, Princeton, NJ: Van Nostrand Co.
- Mills, C. Wright (1959), *The Sociological Imagination*, New York: Oxford University Press.
- Sampson, Robert (2012), *Great American City: Chicago and the Enduring Neighborhood Effect*, University of Chicago Press.
- Raub, Werner, and Thomas Voss (2017), “Micro-Macro Models in Sociology: Antecedents of Coleman’s Diagram”, in Benn Jann and Wojtek Przepiorka (eds), *Social Dilemmas, Institutions, and the Evolution of Cooperation*, Berlin, De Gruyter: 11–36.
- Raub, Werner, Vincent Buskens and Marcel A.L.M. van Assen (2011), “Micro-macro links and microfoundations in sociology”, *The Journal of Mathematical Sociology*, **35**, 1–25.
- Schelling, Thomas C. (1978), *Micromotives and Macrobehavior*, London: W. W. Norton.
- Swedberg, R. (2016), “Can you visualize theory? On the use of visual thinking in theory pictures, theorizing diagrams, and visual sketches”, *Sociological Theory*, **34** (3), 250–75.
- Thornton Patricia H., William Ocasio and Michael Lounsbury (2012), *The Institutional Logics Perspective. A New Approach to Culture, Structure, and Process*, Oxford: Oxford University Press.
- Turner, Charles (2010), *Investigating Sociological Theory*, London, UK: SAGE.

- Udén, Lars (2001), *Methodological Individualism: Background, History, and Meaning*. London: Routledge.
- Ylikoski, Petri (2011), “Social mechanisms and explanatory relevance”, in P. Demeulenaere (ed.), *Analytical Sociology and Social Mechanisms*, Cambridge: Cambridge University Press, 154–72.
- Ylikoski, Petri (2012), “Micro, Macro, and Mechanisms”, *The Oxford Handbook of Philosophy of the Social Sciences* (edited by H. Kincaid), New York; Oxford University Press, 21–45.
- Ylikoski, Petri (2013), “Causal and constitutive explanation compared”, *Erkenntnis*, **78**, 277–97.
- Ylikoski, Petri (2014), “Rethinking Micro-Macro Relationships”, in Finn Collin and Julie Zahle (eds), *Rethinking the Individualism-Holism Debate. Essays in Philosophy of Social Science*, Dordrecht: Springer, 117–35.
- Ylikoski, Petri (2017), “Methodological Individualism”, in Lee McIntyre and Alexander Rosenberg (eds), *Routledge Companion to Philosophy of Social Science*, New York: Routledge.