The Coleman Diagram, Small *N* Inquiry and Ethnographic Causality

Peter Abell (London School of Economics) P.Abell@lse.ac.uk

Ofer Engel (De Montfort University) Ofer.Engel@dmu.ac.uk

Abstract: Certain key ideas have served to stimulate thought about the nature of sociological inquiry by presenting frameworks which are adjustable to the varying demands of different sorts of data. They provide, as it were, heuristics which enable a systematic debate about the optimal simplifications that are enjoined by varying types of data, notably what are sometimes labelled as quantitative and qualitative. In so doing they facilitate a variety of analyses, generate hypotheses and eventually lead to discoveries that transcend inherited boundaries. Our claim is that the Coleman diagram, though largely interpreted within a quantitative tradition, is one such framework for formalizing qualitative methods and for understanding the interplay between different types of data in the social sciences.

Keywords: Coleman Diagram, Qualitative Methods, Quantitative Methods, Causality, Bayesian Narratives

Social scientists frequently seek to analyse groups, communities, organisations and other collective entities, even total societies, as "objects" or units of inquiry. In so doing they document these entities in terms of their properties (e.g. group cohesion), their relationships to each other (e.g. inter group competition) and the causal mechanisms, which bring about changes in their properties and relationships. Let us call this the macro level of causal analysis. For many decades issues have been debated which draw connections between this sort of analysis and a focus upon the individual units within the collective entities, their properties (e.g. gender), their relationships to each other (e.g. interpersonal trust) and the causal mechanisms which drive them. Let us call this the micro level of analysis. In the general context of the relationship between the micro and macro the so-called Coleman Diagram (or boat, see Figure 1) has gained a notable reputation where the arrows are often, though not invariably, interpreted as causal connections of one sort or another, each of which may imply a connecting social mechanism (Hedstrom, 2005) either observed or theoretically postulated[1]. The "boat shape" is used to imply passage of time from left to right. Furthermore, repeated diagrams whereby the exogenous macro cause is the outcome/effect of a previous cycle may be conceived[2].

Figure 1: Coleman's diagram (about here)

Coleman was insistent that the explanatory objective of the sociologist is always to explain the macro (system) outcome or "social organisation" at the top right hand corner of his diagram, and not the individual level (i.e. the bottom right hand corner). Coleman is, thus, clear that the process of discovery must be directed at the macro level. Furthermore, "the minimal basis for a social system is two (micro) actors each having control over resources of interest to the other" (Coleman, 1990; p29). This then implies that the micro level is usually concerned with interdependent (i.e. interactive) micro actions and, though Coleman introduces a rather complex picture of interdependencies, it is perhaps helpful initially to distinguish between strategic and parametric interdependence. The former implying that the actor faces others as a reactive environment and the latter as a fixed environment. Strategic interactions do, of course invite game theory as the most powerful theoretical framework (e.g. Diekman, 1985). Either way, Coleman was a very early advocate of, "structural research which will represent a truly sociological methodology". In this sense he was somewhat ahead of his time and it is only with the subsequent flowering of social network theory that the "sociological methodology", as he conceived it, has begun to bear fruit [3].

Clearly Coleman's diagram invites rather complicated research designs, which is the issue we wish to address in this short paper. How can it direct and fashion research design? As he observes "... it is one thing to trace the development of social organisation in a particular instance, as a historian might do, and quite another to develop generalisations about such processes". We shall argue that as a consequence of the inherent complexity in research designs sociologists may often have to rely upon what we term ethnographic (i.e. small N) causality rather than statistical models of causality based upon frequencies (large N) and covariance (Abell, 2009).

The Coleman diagram generates debates about reduction, methodological individualism, top down and bottom up causality, in contrast to emergence and social wholes and collective causality. Reduction usually implies that any concept deployed at the macro level can be "derived from" or "reduced to" concepts at the micro level, implying that the macro is defined in terms of the micro. This applies equally to the cause, the effect and to the connecting causal mechanism at the macro level. Emergence, on the other hand, denies this to be the case, implying then that macro concepts are, in some sense, sui generis. These issues may be extended to more than two levels when it is then, for example, appropriate to speak of the micro, the meso and the macro levels.

Particularly problematic are issues of causality. Can causal relations and mechanisms at the macro level (arrow 4) always be reduced to micro causality (i.e. the conjunction of arrows 1, 2 and 3)? Or, what amounts to the same thing, can macro causal mechanisms be faithfully constructed from micro causal connections? Coleman (1990, p5) argues that "no assumption is made that explanation of

systemic ("i.e. macro") behaviour consists of nothing more than individual action and orientation taken in aggregate. The interaction among individuals is seen to result in emergent phenomena at the system level, that is phenomena that is neither intended nor predicted....". Coleman, thus, tends to construe emergent macro outcomes as the unintended outcomes of micro actions, presumably also not predicted by the actors involved[4]. He was also concerned to place sociological inquiry within a framework of policy recommendation. Any such recommendation implies prediction which in turn implies causal explanation which usually then implies the detection of generalisations based upon comparative studies. It is therefore useful to see how his diagram relates to such complex ambitions by directing research appropriately.

At the micro level it is perhaps rather straight forward to interpret the causal connection as taking the form: micro cause(s) →individual actions/interactions→ micro outcome(s). That is to say, it is individual actions (or forbearances) and interactions which drive things along, connecting causes to outcomes. Agency provides "the mechanism through which the focal independent variable is able to influence the dependent variable of interest" (Baron and Kenny, 1986). There has been much recent analysis of causality in terms of what are called causal powers. It is then tempting to adopt an entirely parallel picture at the macro level where group actions/interaction (whatever this might mean) becomes the focus of attention. Thus, we would have: macro cause(s) brings about group actions/interactions which brings about macro outcome(s). This simple parallel, however, hides a multitude of possible problems. Indeed, If full reduction is at all feasible then the top arrow (4) as an independent causal relation falls away merely standing as a shorthand for the causal route through the micro level and back up to the macro. If so, how do the vertical (causal?) connections fit into this framework? Indeed, whilst the macro-micro link is pretty un-controversially interpreted as causal, the micro to macro link, as Coleman observed, is not necessarily so easily dispatched, sometimes it involves mere aggregation of micro outcomes (when full attention must be paid to issues of ecological correlation, (King, et al., 2004)), other times it is suggested as causal, implying a contingently independent way of characterising the macro effect. The Coleman diagram systematically directs researchers to engage with these issues at the very outset of their research. It provides a map of problems that must be addressed from the start. Inspection of the papers in many of the journals suggests that Coleman's implicit injunctions are not always taken to heart.

Coleman urged that micro and macro units of analysis are characteristically interdependent and it is widely acknowledged that drawing samples of units as if they are independent, as much survey research does, is likely to lead to misleading results. This implies that we have to take account of structures (networks) of units, at all levels in the context of the causal inferences in his diagram. This,

as we shall see, introduces demanding requirements in order to enable large N statistical treatment as reliable.

Given the demanding requirements and assumptions of large N multilevel studies, which furthermore characteristically involve networks at both levels, it is unlikely that many of the multilevel problems, conforming with the Coleman diagram, which social scientists may wish to address can easily be pursued in a statistical manner. A multilevel analysis when only a few or even a single macro unit is available (either pragmatically or in fact) needs to be fashioned. Since the Coleman diagram is often structured around causal relations, this amounts to finding a way of investigating causality in the absence of inter-unit comparison and generalisation (Abell, 2009;Abell and Engel 2018); that is to say, an ethnographic concept of causality. The diagram carries the virtue that it enjoins the researcher to address issues of large versus small N from a principled standpoint not merely as a personal disposition.

The paper develops as follows; first, the general framework of multilevel networks is introduced; second, Bayesian Narratives will be briefly introduced: third, ethnographic causality is explored as a possible solution to intractable complexities of statistical based causality.

A General Framework, Multilevel Networks

The statistical hierarchical linear model, where units of analysis at both the micro and macro levels (e. g. individuals and groups) can be independently drawn from (often hypothetical) populations, is now well developed (Snijders and Bosker,2012),[5] enabling a coherent research interpretation of the Coleman diagram. The model, however, does not capture Coleman's insistence that units at both levels cannot be drawn independently since their properties depend upon the properties of their neighbours as they are each embedded in their respective networks.

A useful abstract framework, which acknowledges non -independence, places the analysis of the relationship between the micro and macro in terms of three types of multiplex networks (Bliemel, McCarthy and Maine, 2014) defined upon vector labelled nodes i.e. multiplex di-graphs or graphs defined upon vector labelled nodes; one at the macro level, one at the micro level and a bipartite graph assigning micro units into macro units (Snijders and Lazega, 2016). A vector labelled nodes multiplex network consists of a fixed set of nodes (usually finite), each of which may carry various measured properties, and sets of directed or undirected relationship types (edges) running between some pairs of nodes which, in turn, may also carry measured weightings. Networks may, of course, be depicted as matrices of one sort or another. The bipartite graph may allow micro-units to be in more than one macro-unit (i.e. multiple group membership). If so then micro-units with a place in

more than one group will contribute to inter-macro unit relations. In addition, relations running between micro level units which are in turn affiliated with distinct macro-units, contribute to inter-macro unit relations (see Figure 2). Finally, macro-units may be related by "emergent" macro level relations (e.g. group relations not derivable from inter-individual relations).

Figure 2: micro-network (left) and the corresponding macro network (right) (about here)

An example may help to clarify issues. The micro level may comprise a selected group of individuals (i.e. micro nodes), each of whom carries the properties of both weighted group identity and gender (vector labelled nodes of length 2). Some or all pairs of individuals may, to a degree, interact with each other (say an ordinal symmetric relation) and trust each other (say a directed binary relation). The members may exclusively belong to the selected group or to more than one group with an associated strength of identity. At the macro level the nodes are groups, each carrying the property of group cohesion and group size (again vector of length 2) whilst some or all pairs of groups are in inter-group competitive and inter-group domination relationships.

Transferring this complexity to the Coleman diagram, the research objective is to explain either the incidence of or change in group cohesion (the top right hand corner of the Coleman diagram) where cohesion is a property of the group embedded in a macro network (Figure 3). The other three points in the Coleman diagram now also become multiplex, vector labelled node networks. Thus the diagram now has vector labelled node networks at each corner. This very general picture then invites queries as to how aspects of networks causally create modified networks at both the micro and macro levels. This complex picture inevitably arises if we embrace Coleman's insistence that sociology is a structural science. Furthermore, Coleman suggests that if we fail to follow this then sociological knowledge will not become cumulative. We are far from certain that this is correct but it does offer food for thought and his diagram so formulated promotes a general heuristic in formulating sociological research.

Figure 3: Networks and the Coleman diagram (about here)

Within this framework and the example, potential causal relationships at both levels can run between:

> (1) Distinct properties of the nodes, including their structural location in their respective networks e.g. in-degree, centrality etc.in a network (i.e. in network terms, monadic analysis). For example, at the macro-level, the causal inference could be between group size and group cohesion measured across groups. Whilst at the micro level it could be between individual identity with the group and gender across the individuals in the selected (sampled) group.

(2) Distinct relation types between pairs of nodes, (i.e. in network terms, dyadic analysis). For example, at the macro level the inference could be between inter-group dominance and inter-group competition across groups in the macro networks. At the micro level it could be the impact of inter-personal interaction upon inter-personal trust.

(3) Distinct properties of the nodes and the relationship between pairs of nodes, (mixed monadic dyadic analysis). For example, at the macro level the inference could be between a group's cohesion and inter-group competition and at the micro level the impact of gender upon inter-personal trust.

Both levels allow for local measures characterising each node's structural location in terms of their relationships to other nodes carrying properties (e.g. trust in-degree/out degree weighted by identity of structural neighbours) and global measures characterising the whole network (e.g. network trust density perhaps weighted by identity). Given this picture it is important to observe that macro units (like groups) can be said to possess both global-micro properties derived from their individual level networks and macro-local measure in virtue of the position of the group in the macro network.

It is also important to be clear about what constitutes the population of macro level units. They could be networks of relations on group nodes, then the population from which the networks are drawn would all be networks of this sort. But in practice given the constraints of research and the difficulties in defining such populations then the focus will likely be upon some groups in a particular group level network. This might be designated as a meso level where the units of analysis will be specific groups with external relations to other, non-observed, groups in the network. By taking the cohesion of groups as the macro outcome to be explained in the previous example, the groups (with of course have an internal network structure on individual nodes) will be treated as nodes with a node specific pattern of external relation types to other groups.

The statistical (large N) approach, adopting this framework often depends upon two stage sampling – namely a sample of macro units and of micro units each of which is nested in one or more of the macro units (groups). These samples, in turn, depend upon a specification of appropriate populations from which they are drawn. Furthermore, since the units at both levels are located in networks they cannot be independently sampled necessitating specialised statistical inference procedures (Snijders and Lazega, 2016) appropriate to networks. The statistical procedures for analysing this sort of complexity are in their infancy but developing rapidly and should be fully embraced where empirically feasible[6]. However, the data requirements and assumptions are so demanding that it is difficult to see how many of the macro outcomes we might be interested to explain causally can be approached in a fully-fledged statistical manner. Observing samples of

connected macro and micro units, each drawn from defined populations, is clearly daunting. Rather a small number or even a single case of the macro unit is a more likely focus. This implies that we will study a few macro units and perhaps seek to achieve a meta-analysis across other similar but not identical studies whilst retaining Coleman's ambition to explain macro outcomes. A limited sort of statistical analysis can be achieved with only a handful of cases/units using, for instance, Fisher's (1958) method, but any causal inference remains hazardous. Coleman (1990) himself acknowledges these issues both in his analysis of the classical Weberian thesis about Calvinism and Capitalism and in the early "qualitative" chapters of his monumental book. However, the nature of causal inference in these "qualitative" endeavours remains rather obscure. How can we address the complexity of multilevel network analyses where units are not independently sampled and where either data only upon a limited number of macro units is available or comparators are scarce? In either eventuality this rules out systematic comparison and statistical generalisation each of which are the standard ingredients of any causal inference.

Bayesian Narratives and Small *N* Multi-level Analysis.

To express the problem succinctly, we need to furnish a method of causal inference, charting the role which social actions and interactions play, that depends neither upon systematic comparison nor statistical generalisation across cases. Bayesian Narratives claim to provide such a method (Abell, 2007, 2009a, 2009b, Abell and Engel 2018). A narrative comprises a time ordered di-graph where a chronology of nodes represents causal conditions, intervening actions and outcomes and the edges causality (i.e. consistent with causes→ actions /interactions→effects. The Coleman diagram when matched with narratives can then be depicted with narratives lying on causal links 2 and 4 providing the connective mechanisms, the former describing the interactions between groups and the latter between individuals within groups (Figure 4). Thus, two types of narrative provide, as it were, the connecting causal mechanisms, respectively at the macro and micro levels. The arrows 1 and 3 are handled differently –see below.

Figure 4: Inserting narratives into the Coleman diagram (about here)

The question still remains as to how the causal relations (edges) are to be studied in the absence of frequent comparative cases? Ethnographic causality provides one possible route which introduces the idea of a Bayesian narrative. Before however exploring the conception of ethnographic causality it is important to take a view on the nature of arrow 4 in Coleman's diagram, namely upon macro causality.

Macro- Causality?

Do macro causal effects (arrow 4) exist, or can they always, at least in principle if not necessarily in practice, be reduced to the conjunction of arrows 1, 2 and 3 in Coleman's diagram? This is of course a thorny issue in the history of sociology at least since Durkheim's time. If we stand by the assumption that causal links must involve intervening mechanisms, featuring actions and interactions (i.e. driving mechanisms), then to assert the existence of independent macro causes seems to necessitate a concept of irreducible collective action[7]. This would imply that the narrative connecting the exogenous macro- structural distribution to its modification could only be constructed in terms of collective actors embedded in the macro networks. Reductionists would, however, say that when collective actions are correctly conceived they ultimately imply actions by individuals in the collective (group) taken in recognition of and on behalf of the group. Such statements seem to imply that the individual actions are, at least partially caused by the collective level (arrow 1). Although we favour in principle reduction we do not want to take a definitive position on this issue here.

In the context of the Coleman diagram it is worth noting that the causal connection between the exogenous macro cause through arrows 1,2 and 3 to the macro effect/outcome comprises a complex intervening mechanism running between the macro variables which are also, in addition, directly connected by arrow 4.[8] The now standard way of thinking about the impact of direct causes, in observational studies, is due to Pearl (2009) and his concept of causality derivative of a-cyclic directed graphs (DAGS). By fixing the value of the intervening variable (in Pearl's analysis, by deleting all the causal arrows incident into the intervening variable) can surrender an estimate if the direct effect, here the macro causal effect, if it exists. Thus, the emergent standpoint requires that there is no such reduction available that eradicates the direct macro causal connection. This procedure, as we shall see, has implications for the ethnographic causality.

Ethnographic (Small-*N*) Causality

Causal connections between events (in the current context between multiplex, vector labelled node networks) are to be derived from ethnographic investigation of those involved in the generation of the causal link which produces the transformation (arrow 2 in Figure 1). It is essential to acknowledge that ethnographic studies commence with analyses of actions and interactions, not the events they connect. The latter then empirically arise from the former[9]. This may be envisaged as examining the actual arrows in the Coleman diagram.

Careful in depth interviewing can, with due cautions, elicit statements about actions and interactions of the general form (Abell and Engel, 2018):

"I did X_o because of X_c to realise Y" (a subjective causal statement);

"I would not have done (forborne to do) X_o if X_c had not been the case" (a subjective counterfactual statement);

Where X_c, X_o and Y are characteristically conjunctions of events, states or other actions expressed in natural language of those observed [10]. The beauty of such statements is that, if they can be regarded as reasonably credible, we are put in possession of evidence for both the causal connection and counterfactual for the same unit of analysis (Individual). This of course contrasts with statistical analysis where both cannot be observed for the same unit and various comparative procedures have to be resorted to making a strong assumption of across unit homogeneity. It should be emphasised, further, that the claim for a causal link does not derive from the conditional probability connecting the causal events as, for instance in Bayesian nets. Although a causal connection with a high posterior probability of existing based upon credible statements (i.e. a case study) may provide evidence for computing such probabilities.

Furthermore, counter- potential statements of the general form;

"I could have forborne to do X_o in X_c", and

"I could have done X_o in the absence of X_c"

can also in principle be elicited, which effectively maintain the voluntary nature of causality generated by human agency. Some ethnographers are reluctant to evoke any concept of social causality precisely on the grounds that they believe intentional actions cannot be causally determined and this reservation also often extends to a rejection of statistical causality based upon generalisation (covariance) and inter-unit comparison. Subjective counter - potentials, if deemed credible, can alleviate these reservations.

First, person subjective causal, counterfactual and counter-potential statements can also be complemented by third person renditions- she did X_o etc. In addition, first and third person plural (we, they) statements may apparently allow for the idea of collective causality (arrow 4 in the Coleman diagram). However, the reservations expressed above about collective action should be kept in mind.

All depends, of course upon the credibility of such statements which in turn depends upon the credibility of their proponents as sources of information about their own and others actions/interactions as estimated by ethnographers in their social interaction with them. The elegance of ethnographic inquiry derives from the notion that justified, credible and mutual belief can arise from the social constructions arrived at in social interactions between the ethnographer

and subjects. Ethnographers usually commence inquiry by asking for the meaning of certain actions which is sometimes expressed as determining "what is going on here".

In the present context the focus is, firstly, upon actors embedded in networks at the micro levels and how their experiences (set X_c) causes them to act in the way they do. Secondly, the focus switches to the consequences of the actions /interactions (set Y).

It is imperative to contrast the methodology involved in ethnographic inferences to that involved in, what we might term, standard social network theory. The latter starts with a given set of units and derives a picture of their binary relations and node properties, usually making use of standard interview schedules. Ethnographic inference, on the other hand, painstakingly constructs the networks on the basis of the actors' elicited causal and counterfactual statements (Abell and Engel, 2018). In the case of set X_c, the causes of a given action X_o are sought in the case of set Y, the effects of a given action is the focus. In the case of small groups then this may involve all the group members but for larger groups the ethnographer (s) may have to carefully select members and resort to standard ethnographic concepts of convergence in order to resolve when to terminate the analysis.

Characteristically, a particular actor's set X_c may, at the micro level, refer to others' actions and their properties and to their own ego network position. Reference may also be to properties of the group and to the group's location in the macro network. In this sense the multi-level analysis can be derived from the subjective causal and counterfactual statements. Set Y can be similarly be used to fabricate the modified micro network. The construction of causes and effects in this manner require the difficult to acquire, skills of what we might term a local historian. The sets X_c and Y, it should be noted, are conjunctive causes and effects. Being postulated in a particular case they cannot logically be alternatives. Thus, the constructed narrative is an "and graph" unlike the "or graph" models in statistical structural modelling.

Returning to the much debated causal connection at the macro-level (arrow4); firstly subjective statements in the plural may be used to imply collective actions but with all reservations expressed above. If actors specify, in the context of set X_c, macro properties of macro units then these will count as contributing to arrow 1 in Coleman's diagram. Similarly set Y may refer to macro outcomes -"I/she did X_o to help generate group cohesion Y". However, more likely, set Y will not refer to macro outcomes but to changes contributing to the modified micro structural distribution (bottom right hand corner). Then the macro- outcome is some summary aggregate measure of this structural distribution (e.g. network density weighted by individual identity). Whether this should be regarded

as causal is controversial. Since the mapping from states of the structural distribution to the macro state will be many to one the relationship between a particular micro-state and the macro state is only probabilistic and not, therefore, like a one to one conventional definition some regard it as causal.

Conclusions

The Coleman diagram provides an indispensable guide to the construction of genuine causal analyses of macro level outcomes. It has traditionally invited statistical (large N) analyses which are probably now best interpreted in terms of (linear) network hierarchical models which run faithful to Coleman's structural perspective. Such models should always be the first choice of sociologists, but given the interdependencies between entities in the data, often impose such demanding requirements and challenging assumptions upon the investigator to render them impractical. Moreover, when studies require the specification of detailed causal conditions, the number of comparative cases dwindles to the point where statistical techniques become inappropriate. Causal analysis has always been exclusively associated with a generalising, comparative large N perspective which has led many small N "qualitative researcher" to dispense with the concept altogether. However, Bayesian Narrative Analysis (which should not be confused with Bayesian Networks) begins to open up a systematic way of inferring causality based upon subjective causal, counterfactual and counter potential evidence, where any limited generalisation across cases is posterior to causal explanation, not a presumption of explanation[11].

Foot-Notes

(1) Coleman's use of the diagram was anticipated by a number of authors, though it has become indelibly associated with his name (Raub and Voss, 2016).

(2) We have labelled the diagram (unlike Coleman) in terms of causes and effects as this paper will concentrate upon causal relations. It should be understood that when thinking within a statistical framework all matters of the distribution of residuals and unobserved covariates producing spurious effects have been solved. Thus, the inference from co-variation to causality is as unproblematic as is feasible.

(3) The assumption of independence derives from the stable unit treatment value assumption (SUTVA) underlying causal inference in most statistical (large N) studies. This assumption requires that the causal outcome of a unit of analysis is independent of all other units and the method of treating the unit. In the context of this paper it is noteworthy that some social scientists refer to this assumption as "no macro effect".

(4) Presumably Coleman means a non-intended/predicted collective outcome. This may be an aggregate of the outcome of individual actors, when the aggregate/collective concept is reducible even though it was not predicted/intended by the actor. The fact that it was not intended/predicted does not invalidate a reductive stance. The argument against reduction and for emergence has to be found elsewhere (see below).

(5) The Coleman diagram was proposed in period when the role of a macro level variable varying across groups would likely be interpreted by a fixed effect regression equation sometimes called a contextual effect equation. This formulation, however, assumes that the regression on the micro level variables do not vary across groups. We now re-formulate the model as the hierarchical linear model to allow for both coefficients which vary across groups and errors at the micro and macro levels. Thus, with standard precautions and adopting the vocabulary of cause and effect to depict statistical relations the hierarchical linear model allows that a macro (group) level, (cause) (Coleman top left) has both a direct effect and an interactive effect with a exogenous micro cause (Coleman bottom left) on the micro effect (Coleman bottom right) (Sjniders and Bosker, 2004). The approach is extendable to many micro and macro causes. Note that this formulation of arrow 1 in the Coleman diagram does not run a causal link between the macro cause and the micro cause. Rather in the standard interpretation of the macro cause impacts the micro effect directly and, in addition, interacts with the micro cause. In this latter effect the macro variable impacts the mechanism (i.e. actions and interactions) connecting the micro cause to micro effect (arrow 2). We shall see below this interpretation has virtues in respect of ethnographic causality.

(6) A number of statistical techniques for analysing networks as either independent (causal) or dependent (effect) variables are available. However, models for both at the same time are not fully developed.

(7) The term collective action is used in a variety ways in the literature. The powerful literature on game theory usually studies how individual interests can generate a collective (Nash) equilibrium. But note there is an implicit reductionist viewpoint being expressed here. More generally, the term is used in propositions of the form "the group did X_o". As a pragmatic contribution such statements are indispensable but can they in principle always be reduced to statements about individuals doing things in the name of the group. Note, if so, then this looks like an arrow 1 in Coleman's diagram. (See below).

(8) Since the macro outcome might be over-determined, in the sense that alternative causes are sufficient but not necessary for a causal connection, then some of these links may be reducible and

others not. Issues of non-observed confounding spurious relations at the macro level might also arise. Then Pearl's concept (2009) of the backdoor will be a consideration.

(9) A large N (statistical) study will characteristically find a co-variation between events and then either theoretically or by observation postulate the mechanism that purportedly connects them.

(10) Clearly alternative locutions are possible.

(11) Some readers may have seen a parallel between the issues raised by the Coleman diagram and the fraught debate about group selection in evolutionary theory. Indeed, some authors use the term co-evolution to describe the implied dynamics connecting the micro and macro in Coleman's diagram. However, dynamic processes are not necessarily the same as evolutionary selection and loose parallels are dangerous. It is not at all clear from the Coleman diagram what the selective units would be.

References

Abbott, A. (2004). *Methods of Discovery: Heuristics for the Social Sciences*. New York, N.Y.: Norton & Company

Abell, P. (1987). *The Syntax of Social life: The Theory and Method of Comparative Narrative,* Oxford, Oxford University Press.

Abell, P. (2007) "Narratives, Bayesian Narratives and Narrative Action", *Sociologica,* 3, 1-20.

Abell, P. (2009a). "A Case for Cases: Comparative Narratives in Sociological Explanation", *Sociological Methods and Research*, 32, 1-33.

Abell, P. (2009b). "History, Case Studies, Statistics and Causal Inference", *European Sociological Review,*25, 561-569.

Abell, P. and Engel, O. (2018). "Subjective Causality and Counterfactuals in the Social Sciences: Towards an Ethnographic Causality?" *Sociological Methods and Research*, (submitted, being revised and resubmitted).

Auyang, S. Y., (1998), *Foundations of Complex System Theories in Economics, Evolutionary Biology and Statistical Physics,* Cambridge, Cambridge University Press.

Baron, R. M. and D.A. Kenny (1986). "The Moderator-mediator Variable Distinction in Social Psychological Research: Conceptual Strategic and Statistical Considerations", *Journal of Social Psychology,* 51, 1173-1182.

Bliemel, M.J., McCarthy, I.P. and Maine, E., (2014). An integrated approach to studying multiplexity in entrepreneurial networks. *Entrepreneurship Research Journal*, 4(4), pp.367-402.

Coleman, J.S. (1990). *Foundations of Social Theory,* Cambridge Mass, Belknap Press.

Dennett, D. C. (2013). *Intuition pumps and other tools for thinking*. WW Norton & Company.

Diekman, A. (1985). "Volunteer's dilemma" *Journal of Conflict Resolution,* 29(4), 605-610.

Fisher. R.A. (1958). *Statistical Methods for Research Workers, London,* Hafner Press.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological review*, 98(2), 254.

Hedstrom, P. (2005). *Dissecting the Social. On the Principles of Analytical Sociology,* Cambridge, Cambridge University Press.

Kenny, D.A., Kashy, D.A., Cook, W.L. and Simpson, J.A., 2006. *Dyadic data analysis (Methodology in the social sciences)*. New York, NY: Guilford.

King,G., Rosen, O. and M.A. Tanner (2004). *Ecological Inference, New Methodological Strategies.* Cambridge, Cambridge University Press.

Pearl, J. (2009). Causality: Models, reasoning and Inference, Cambridge, Cambridge University Press.

Raub, W. and T. Voss (2017). "Micro-Macro Models in Sociology: Antecedents of Coleman's Diagram". In *Social Dilemmas, Institutions and the Evolution of Cooperation, Festchrift for Andreas Diekmann,* eds. Jann, B. and W. Przepiorka*,* Berlin, De Gruyter.

Schneider, L. (Ed.). (1967). *The Scottish Moralists on Human Nature and Society*. Chicago, IL: Phoenix

Snijders, T.A.B. and R.J. Bosker (2012). *Multilevel Analysis,* London, Sage.

Snijders, T.A. B. and E. Lazega, eds. (2016). *Mulilevel Network Analysis for the Social Sciences,* London, Springer.

Wrong, D.H., 1961. The over socialized conception of man in modern sociology. *American sociological review*, pp.183-193.

Figures

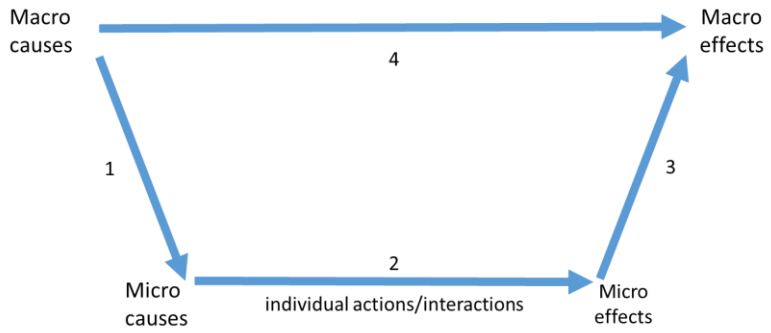

Figure 5: Coleman's diagram

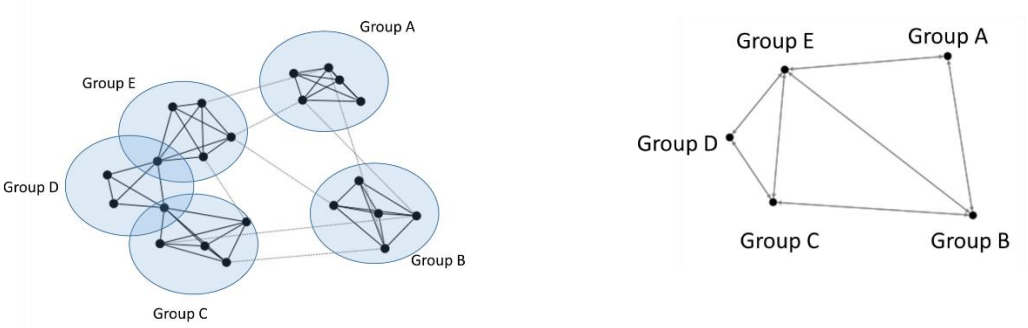

Figure 6: micro-network (left) and the corresponding macro network (right)

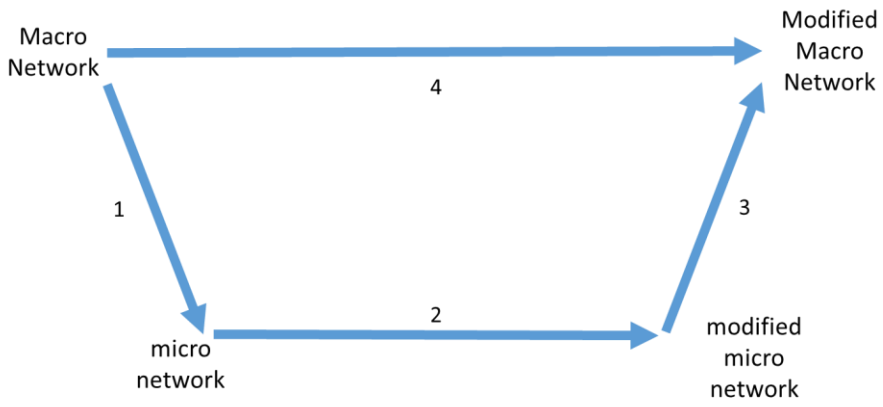

Figure 7: Networks and the Coleman diagram

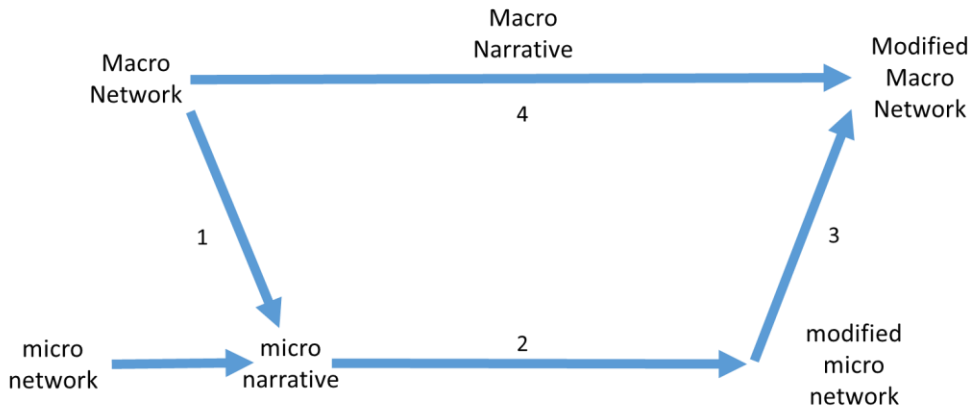

Figure 8: Inserting narratives into the Coleman diagram