Interests, Folk Psychology and the Sociology of Scientific Knowledge

Petri Ylikoski

University of Helsinki

This paper provides a conceptual analysis of the notion of interests as it is used in the social studies of science. After describing the theoretical background behind the Strong Program's adoption of the concept of interest, the paper outlines a reconstruction of the everyday notion of interest and argues that the same notion is used also by the sociologists of scientific knowledge. However, there are a couple of important differences between the everyday use of this notion and the way in which it is used by the sociologists. The sociologists do not use the term in evaluative context and they do not regard interests as purely non-epistemic factors. Finally it is argued that most of the usual critiques of interest explanations, by both philosophers and fellow sociologists, are misguided.

The concept of interest is a part of everyday vocabulary – our folk psychology – that we use when making sense of each other’s behavior. This concept is also widely used in the social sciences both for heuristic and explanatory purposes. For example, in the social studies of science, the interests belong to the explanatory repertoire of many historians of science, philosophers of science, and sociologists of science.¹

The widespread use of ‘interest’ vocabulary does not mean that there is a consensus concerning its proper use. Especially in the social studies of science, explanations put forth in terms of social interests have become controversial. There are disagreements both about the nature and the scope of these explanations. At the centre of these controversies is the so-called Strong Program in the sociology of scientific knowledge (SSK), which is often assumed to advocate some sort of special theory of interests. In this paper I will argue that contrary to a widely held opinion, the Strong Program uses the same folk psychological concept of interest as does everybody else.

I will start my argument by describing the theoretical background behind the Strong Program’s adoption of the concept of interest. Next I will outline the everyday notion of interest and its relation to some other folk psychological concepts. In the third section, I will take a look at how the concept of interest is used in the everyday discourse by scientists. Finally I will show how misguided most common critiques of interest explanations are. The emphasis of this paper will be on conceptual analysis. Thus, I do not intend to present a historical account of the influences or the details of various applications of interest concepts in the social studies of science. Nor will I deal with the uses of interest concepts outside the SSK.² The present aim is instead to de-mystify the use of the interest concept in the social studies of science.

1. The Absence of the ‘Interest Theory’

¹ For example, see Bourdieu 1977; MacKenzie 1981; Whitley 1984; Rudwick 1985; Shapin & Schaffer 1985; Sapp 1987; Latour 1987; LeGrand 1988; Giere 1988; Hull 1988; Fuchs 1992; Biagioli 1993; Kim 1994; Lenoir 1997; Gieryn 1999; Segerstråle 2000.

² For the history of the concept see, Hirschman 1977; Myers 1983.

There is a widely shared view that interest explanations in the social studies of science are most often related to *the interest theory* developed by the members of the Strong Program. (Barnes 1977; Barnes & MacKenzie 1979; Barnes, Bloor & Henry 1996) I claim that this is an illusion: almost everyone uses interests in their accounts of science, but only the members of the Strong Program have elevated this concept to the status of a central theoretical term. Nevertheless, adherents of the Strong Program use the same folk psychological concept of interest. To substantiate this claim, I will first take a brief look at the early work of Barry Barnes.

The central epistemological idea and the justification for the possibility of the SSK in Barnes' early writings is the thesis of *instrumentalism*. This thesis claims that knowledge-producing activities are goal-directed processes in which agents use their earlier knowledge as a resource to create new knowledge and to achieve their other goals. Agents are thus seen as innovative and selective users of their cultural inheritance. Barnes contrasts this conception of knowledge production with the *contemplative* account of knowledge. According to this account, disinterested individuals produce knowledge either by passive perception or by contemplation. Here knowledge is taken to reflect or correspond to reality in a very straightforward manner. In this conception, only the mistakes and invalid descriptions of reality are taken to be due to the influence of social or personal interests disturbing the normal process of perception. (Barnes 1977, 1-2.) By arguing for instrumentalism, Barnes tries to establish that: 1) both the production and the maintenance of knowledge are essentially social processes, 2) knowledge production is an active and goal-directed process, and 3) the relation between knowledge production and certain goals and interests is constitutive rather than external. If these theses are granted, the SSK project is justified.

According to Barnes, the connection between knowledge and interests is neither internal nor logical. Rather, he maintains that interests inspire the construction of knowledge out of the available cultural resources in ways that are specific to particular social and cultural contexts. He also notes that knowledge does not appear and disappear as various kinds of interests wax and wane. Instead, knowledge often continues in an intellectual tradition, as a resource to be deployed

in the furtherance of whatever interests are institutionally predominant. In this way, yesterday's 'ideology' can sometimes be transformed imperceptibly into today's 'science'. The validity or later usability of a certain piece of knowledge is therefore not tied to the interests that originally created it. This is an important point. The fact that the choice or the development of a belief, idea or theory has been influenced by some interest, or that it has served some interest, has no automatic consequence concerning its validity or its later applicability for some other purposes. (Barnes 1977, 41-42.)

Instrumentalism is a general epistemological position, not a substantial sociological theory about science. The main sociological implication of instrumentalism is that scientific activity can be studied just like any other social activity. Barnes' discussion of instrumentalism takes place at the same level of generality as does Jürgen Habermas' theory of knowledge constituting interests, which served as an inspiration for Barnes (Habermas 1972; Barnes 1977). Neither of these theories is to be used directly in the explanation of specific scientific episodes. They both share the general thesis that scientists have an interest in 'prediction and control'. This thesis does not say anything about the particular scientists or their goals. In other words, it is a claim about what is *common* to all knowledge production. Theories on this level of generality are not able to account for the details of the concrete episodes in the history of science. To be able to do that, instrumentalism needs to be supplemented with a theory of specific goals and interests that drive scientists and other relevant social actors (Yearley 1982, 359-361).

So, what is the substantial sociological theory about social interests that could provide explanations for the details of specific episodes in the history of science? The surprising answer is that advocates of the SKK do not have one. They have discussed various suggestions towards a theory of goals and interests, but no single theoretical model has achieved a dominant status.

For example, the analyses of historical case studies in terms of class interests were made in the latter part of the 1970's, the two most notable examples being Donald MacKenzie's work on British statistics (MacKenzie 1981; MacKenzie & Barnes 1979) and Steven Shapin's work on Edinburgh

phrenology (Shapin 1979). These two works, together with Barnes' (1977) critical review of various Marxist theories of class interests, clearly show that the theories of class were seriously considered as being a theoretical resource for understanding science and its relation to the surrounding society. However, class interests were neither claimed to be the only, or even the most relevant, interests in explaining scientific episodes, nor to be an essential part of the Strong Program.

Subsequent developments have shown that the class interests as determinants of scientific development were only a temporary hypothesis. This hypothesis was dropped because the case studies showed that in explaining scientific episodes in detail, class interests are a relatively unfruitful explanatory resource. (Ylikoski 2001, Chapter 6) in fact, the references to class interests vanished completely from the SSK literature after the beginning of the 1980's. But this does not mean that the concept of interest was also dropped. For example, Donald MacKenzie and Steven Shapin still use the concept of interest in their later works, but without references to social classes (Shapin & Schaffer 1985; MacKenzie 1990).

Contrary to a common stereotype, the early Strong Program did not have any special preference for the externalist explanations of scientific episodes. For example, Barnes wrote in 1977 about his and MacKenzie's case study of the controversy between the biometricians and the Mendelians:

In the case of scientific controversy, technical factors and esoteric professional interests must always be looked to first as a source of explanation. But in the present instance no sufficient basis for the dispute can thereby be found. The opposed forces did not have access to different kinds of evidence with conflicting implications, nor were differences in their training, and in the skills and competences they possessed and valued, of such magnitude and significance as to account for their different theoretical perspectives. Nothing in the esoteric scientific context satisfactorily accounts for the controversy. Nor there was any technical reason why the disputants should not have agreed to await further evidence, or accepted that both their accounts might have had merit and applied to different kinds of evolutionary change (Barnes 1977, 59).

This passage clearly shows that the intention was not to explain everything in terms of class interests or some other 'external' factors. On the contrary, the passage suggests that Barnes and MacKenzie clearly prefer an internalist explanation for the controversy. Only when explanations in

these terms were found wanting did the authors turn to the wider social influences. For SSK, the extent to which the history of science displays external influences is purely a matter-of-fact issue. (For a discussion of the internalism/externalism issue in history of science, see Shapin 1992.)

After the demise of the theories of class interest, no explicit sociological theories of interests have been advanced. The ‘goals and interests’ (Barnes 1982; Barnes, Bloor & Henry 1996) have remained an abstract formula that has been filled in various ways in different case studies, depending on the details of the case. Furthermore, theoretical discussion about the *kinds* of interests that influence science has yet to emerge. The most popular scheme of analysis has been the model of professional interests of scientists, but it has not received much theoretical attention. (Ylikoski 2001, Chapter 6). These observations suggest that there is no sociological theory about interests, only a general epistemological thesis that should be acceptable to everyone.

2. What Are Interests?

In ordinary language, the notion of interest is used in various ways. It can, for example, be used to refer to a share in a business venture, to a legal concern or right to property, to money paid for the use of borrowed money, and to a hobby. These usages are peripheral from my point of view. Interest can also refer to things or topics that draw one’s attention, concern, or pique one’s curiosity. The attitude can be expressed by phrases such as ‘being interested in’ and ‘being interesting’. One can also make another person interested in some issue or item, which means that one succeeds in attracting that person’s curiosity or succeeds in enlisting a person in a given cause (Callon & Law 1982; Latour 1987).

The most important use of the notion of ‘interest’ is related to things that are conducive to one’s goals or aims. Things are in a person’s interest, or concerns, because of their relation to the achievement of personal goals. These things are not ‘interesting’ in themselves, they are the *means* to achieve one’s goals or to fulfill one’s wishes. This connection ties interest closely to the cluster

of goal concepts which are at the heart of intentional psychology. Among these concepts are such notions as: aim, desire, end, goal, objective, purpose, and want. There are important differences between these notions, but I will disregard them here, since they behave quite similarly in relation to the notion of interest.

One's goals determine the things that are in one's interest. On the other hand, interests tie the goals to the action environment by picking up those things that are necessary for the realization of these goals or are otherwise conducive to them. If one does not and will not have any desires or aims, then one does not have interests either. Interests are *goal-dependent*: there are no interests without goals or aims. In one sense, one could have interests without *currently* having any goals and desires. In this case the interests refer to things that are conducive to the goals that the agent will or can have in the future. For example, we can speak about the interests of a child. However, interests deriving from possible future goals are not helpful in explaining the agent's *current* behavior, so I will not include this very special sense of interest in the present discussion.

The goal dependence of interests makes them goals themselves. If the achievement of *B* is the (only) means to achieve my goal *A*, and I know this, then the achievement of *B* becomes my goal. These goals formed by this kind of practical reasoning can be called *subgoals*. They derive from having some other goal and they are pursued only because of their connection to this more fundamental goal.

One important conceptual difference between goals and interests concerns knowledge. Although it might be sometimes difficult to say what one exactly wants or what one is aiming at, it is clear that there is a sense in which one has to know what one wants or aims for. It is not *my* goal if I do not know that I have it. One can, for a short period of time forget what one was trying to achieve, but if this forgetfulness endures too long, one simply no longer has that goal. The same goes for beliefs about one's goals. One cannot have completely wrong beliefs about one's aims and desires. Thus, what I believe to be my goals, are my goals.

Interests are different. It is possible that the agent is completely ignorant of the things that are in her interest, but those things are nevertheless in her interest. She can also have false beliefs about her interests. She can believe that certain things are in her interest, when in fact they are not. These states of affairs do not produce conceptual problems. How then can we account for this conceptual difference between goals and interests? The agent's goals and desires are subjective in the sense that they are psychological facts about the agent's attitudes. There might be an important difference between goals and desires to the effect that one cannot change one's desires at will in the sense that one can change one's goals, but both concepts are still subjective in this sense. Nevertheless, *once there is a set of goals, it is an objective matter which things are in the agent's interest*. The things which are conducive to one's goals depends on causal facts about one's action environment, not on one's wishes or desires. Thus, I cannot decide that some things are in my interest unless I change my goals.

Moreover, it is possible to have *conflicting* goals. The pursuit of one aim can thereby hinder or completely prevent the achievement of another. For example, one's short-term desires can be in conflict with one's long-term projects or one's various goals can be incompatible. The agent has to solve this conflict by some decision-making procedure. She might let one interest override the others or make some kind of compromise. In such cases, the external observer can infer the relevant interests from the agent's behavior. Conflicting interests are more troublesome for the agent than for the external observer.

However, the external observer faces problems where *the same action is conducive to two different goals*, since it can be difficult to decide which interest was the decisive factor. The problem is analogous to causal overdetermination. The agent herself does not have to decide about the relative importance of her goals, so this happy situation does not create any problems for her. The external observer's problem is that she cannot justify her claims about the explanatory relevance of the particular interest simply by noting its compatibility with the agent's choices and actions. The agent may well have the goal attributed to her and have the relevant knowledge about the things that are

conducive to the achievement of that goal, but these factors are not sufficient to explain her behavior. In fact, her action might have been due to some other interest. To pick the right explanatory interest, the analyst has to know what other interests the agent has. In addition, the analyst needs to have evidence concerning the cases where these interests are in conflict to the effect that the interest the agent has picked is the overriding one. This makes interest explanations holistic: how an agent's certain interest contributes to her behavior depends on her overall interests and goals. The explanation has to take into account all the relevant goals and interests. Otherwise, it might end up without any counterfactual force.

Interests are goal-dependent, but they do not change every time the agent changes her aims, so they do not have to be *goal-specific*. It is possible that the things that are in the agent's interest remain the same. The same things can be conducive to both one's earlier and one's later aims. Some things, like nutrition, are preconditions for the achievement of almost any goal. One does not even need to know an agent's specific goals to know her interest in adequate nutrition. The same goes for wealth. In a modern society, money is a generalized means of achieving things. People go to work to earn money, but rarely is having money itself their ultimate goal. Rather, they want to earn money as a means to access whatever they happen to want, as money can be converted into (almost) everything. Wealth does not guarantee the achievement of one's goals, but it certainly helps in many cases. As a consequence, most people have at least some interest in having money at their disposal, and this interest remains stable while the agents' specific goals constantly change.

In the SSK, the idea is that a scientist's professional interests have this kind of *stability*. The scientists have an interest in their scientific reputation or credibility and it works much like money does in economic life. One's reputation is a generalized means to achieve whatever one is after in academic science. It does not matter whether one is looking for the truth, power (in an academic community), a good salary, or the Nobel Prize: to be successful in the pursuit of this goal one also has to be successful in the accumulation of scientific capital. One's goals cannot be achieved without it. (Barnes 1985, 45-48; Ylikoski 2001, Chapter 6.)

The *stability* of interests is the main reason for the usefulness of the concept. With the help of interests, one can describe the structural features of the agent's action situation and not just their internal psychological states. For example, the model of cycles of credibility described by Latour and Woolgar (1986) is not about the scientist's motives, but about the means by which they achieve their goals. The model cannot directly explain choices and actions by individual scientists, but it can be used explain the general patterns of behavior among scientists. It also works as a heuristic tool both in the search for and the construction of explanations for particular scientific episodes. The model creates a set of expectations against which more detailed explanatory hypotheses can be contrasted. (Ylikoski 2001, Chapter 6.)

It is often thought that interests are always somehow selfish. This is true only in a trivial manner. One's interests are always related to one's own goals. My interests cannot be related to your goals unless I have adopted the advancement of your goals as my own goal. I call this trivial, because the concept of interest does not set any limits for the kinds of goals an agent can have. In other words, the goals can be as selfish or altruistic as the agent wishes. If my goal is to save humankind, then it might be in my interest to solve the world's hunger problem. It is only essential that my interests derive from *my own* goals. In ordinary parlance, when interests are sometimes set against the common good or altruistic behavior, the reference is actually to interests that derive from selfish goals. Naturally people often have selfish goals, but using the concept of interest does not commit one to the view that *all* human action is selfishly motivated. That position would require a separate argument. Thus, there is a clear conceptual difference between an interest and a self-interest.

Interests are often attributed to collective agents such as groups, institutions, professions, classes, firms, and states. Attributing interests to collective agents is often taken as quite unproblematic: if we can attribute goals to a certain collective, then we can also attribute interests to it. After all, being a person or a human is not a prerequisite for having interests. All kinds of agents can have interests, as long they are agents with goals. However, the attribution of interest to a collective is often ambiguous. When one is speaking, for example, about the interests of a profession, one can

refer to two different things. First, the reference can be to the interests of the members of that profession. In this case, all (or most) members of the profession have similar individual interests. I call this *a shared interest*. In such cases, all members can pursue their (shared) individual interests independently of each other. This is not so when one attributes *a collective interest*. In this case, the members of the profession have an interest as a collective. Although even in this case, the ultimate benefit from the satisfaction of this interest goes to the individual members, whereas the means of attainment of this satisfaction are different. The pursuit of a collective interest presupposes collective action; the members of the group cannot pursue it separately. It is not obvious that the agents can in fact act collectively. For example, a conflict might arise between the interests of the collective and the individuals who are its members, which makes the incidence of collective action improbable. Only if the members can solve the problem of collective action can they act as a collective and pursue their collective interest. Usually this requires some internal organization. Firms and states often have such an organization, but social classes usually do not. Consequently, the attribution of interests to states is less problematic than is the attribution of interests to social classes.

What about the use of interests in explanation? Explanation in terms of interests requires certain assumptions. The first assumption is that the agent in question has at least some aims, goals or desires. I call this the *intentionality* assumption. There are not very strict limits on what the agent can want, but he or she must want something. The second assumption is about *rationality*. Agents should be rational in the sense that they do things that they believe to be conducive to their goals or aims, and if there is more than one way to advance their goals, they will choose the one that they believe to be the most conducive to their goals. The third assumption is about *knowledge*. The agent should be able to collect reliable information about her action environment either by her own observations or by using other social agents as the sources of information. Furthermore, The agent has to use these abilities to acquire beliefs about her action environment. Of course, it is not necessary to assume that the agent can take note of everything or knows everything about her action

environment. The assumptions about knowledge give some space for mistakes and omissions, but do not allow that the agent can be completely wrong about her action environment.

A very important distinction needs to be made between action that simply serves a certain interest and the pursuit of that interest. In the first case, the results of an action are conducive to the interest, but the connection between the action and its serving a certain interest is coincidental. The action was not taken because of that interest. As a consequence, the interest cannot explain the action in such a case. In short, only the pursuit of the interest could explain the action.

As intentional explanations are explanations in terms of an agent's beliefs and desires, it is clear that one cannot explain the agent's action by interests the agent is not aware of. The beliefs of the agent always mediate his or her interests in an action. This makes interests redundant in the sense that when we know what an agent's beliefs and goals are, so interests actually do not add anything causally relevant to the explanation.

However, this redundancy concerns only the explanations of individual actions. The concept has also some other uses that are relevant from the point of explanation. First, as noted, interests are important for the reconstruction of the agent's beliefs and goals. They can work as *an explanation-seeking heuristic*. When we are using the interest perspective as an explanation-seeking heuristic, we are not assuming that the explanatory assumptions of the interest explanation are satisfied. Rather, we are using the idea that the agents use the available resources to advance their goals to make sense of scientific activities. An explanation-seeking heuristic helps in the search for candidates for explanatory factors and even in making sense of empirical material before any explanatory questions have been raised. The value of this kind of heuristic does not depend on whether the search will end with an actual interest explanation. The use of the heuristic can also facilitate research by showing that certain interests are not a relevant explanatory factor in the case under study. In my opinion it is beyond doubt that the interest perspective has proved itself to be useful as a heuristic in the SSK.

Second, not all interest explanations are explanations of individual actions. Most of these explanations are structural and are best interpreted as explanations of typical patterns of behavior, not individual actions (Ylikoski 2001, Chapters 6 and 7). The interest explanations in the SSK are a mixed lot. Here no one single pattern of interest explanation can be found, although they all employ the same basic notion of interest which is articulated in this paper.

3. Interest Talk Among Scientists

As interests are part of the repertoire of our mundane reasoning about social action, it is not a surprise that scientists also use the concept of interest to make sense of each other's behavior. In fact, their use of an interest perspective is comparable to that of many sociologists in terms of its sophistication. This gives us an extra motive to study interest explanations: we need to understand them not only to make sense of what people in science studies are saying, but also to make sense of how scientists themselves make sense of the social world around them. The fact that informants are versed in the vocabulary of interests gives an advantage to a historian or sociologist studying scientific episodes: informants are doing part of the analysis. However, their accounts are to be taken with a grain of salt. Some features in the interest talk of scientists are very important to keep in mind, especially when compared with the sociological use of interest concepts.

In scientific controversies, it is often found that the disputing parties try to explain their disagreements by referring to social interests (and also other 'non-cognitive' influences). In Gilbert & Mulkay's terminology, this use of interests belongs to *the contingent repertoire* of the scientists (Gilbert & Mulkay 1984). This repertoire is used to explain away discrepancies between what is believed by the agent herself to be true or rational and what the other people claim to be the truth. This explanatory setting is familiar from everyday life: an agent tries to explain the difference between her own view and the position of someone else. The agent typically assumes that her own position is well justified. (Why would she hold such a view if she did not?) The existence of

disagreement creates an explanatory problem for the agent: why do others not hold the position that she judges to be the right one? In such a situation, the position held by others is regarded as exceptional or abnormal, that is, as something requiring an explanation.

Many sociologists of scientific knowledge have noticed the *asymmetrical* nature of these explanations. The beliefs that are believed to be true or rational are thought to be natural and not in need of any special explanation, whereas ‘false’ beliefs are accounted for in terms of ‘non-cognitive’ factors such as psychological idiosyncrasies, incompetence, social or career interests. (Gilbert & Mulkay 1984.) However, it seems to be an overstatement that scientists *always* offer asymmetrical explanations when accounting for divergence in opinions. It is also possible that they offer symmetrical analysis, at least for some episodes (Tammi 1999). In such a case, both opinions are seen as reasonable, and the divergence is attributed to an insufficiency of data and inconclusiveness of arguments.

Apart from the cognitive aim of explaining discrepancies in opinion, there is also another motivation for this asymmetrical setting: the wish to discredit the competitor’s positions. The idea is that by revealing the ‘true’ motives or causes of opponents’ views, they can be disregarded as non-serious contributions to the debate (Shapin 1979.) We might call this *the exposing stance*. The true motives are contrasted with the claimed motives in a manner that challenges the reliability and credibility of the opponent. The implicit claim is that the opposition supports its view because it serves its interests rather than on account of its epistemic credentials.

An example of this usage can be found in the discussion of conflicting interests in science. In the context of this discussion, interests are understood as being biasing factors that harmfully interfere with ethically sound scientific conduct, most commonly in the form of financial or personal interests. Financial interests refer to the economic stakes that scientists (or their relatives) might have in the results or interpretations of a scientific study. The much fuzzier class of personal interests covers such personal relationships as being a close relative, having a personal vendetta towards somebody, having an ongoing scientific priority conflict, having a teacher-student

relationship, *etc.* (Resnik 1998). This conception of interest is extremely limited in contrast to the concept used in the SSK.

In the everyday talk of scientists (and philosophers), the term interest is most often used for goals that are non-epistemic. So when they find a sociologist of scientific knowledge discussing social interests in science, they naturally assume that the sociologist is speaking about the influence of the non-epistemic factors in science. Usually it is also wrongly assumed that the sociologists of science are only interested in the non-epistemic aspects of science. However, the sociologists of scientific knowledge have not limited the use of 'interest' in this way. They do not equate social with non-epistemic, but treat it as a general category which includes both epistemic and non-epistemic factors. For them, to say that research choices or evaluations are goal-directed does not mean that they are inadequate, unscientific, or biased (MacKenzie 1981, 219).

This denial has two aspects. First, the sociologists of scientific knowledge claim that the impact of social interests is not always negative. Thus, even completely non-epistemic interests can contribute to the production of knowledge by leading scientists to carefully scrutinize some aspects of nature which might otherwise have remained unexamined, or they can make scientific research possible by making resources available for research. In this way, social interests can also be accelerators and motivators. (Shapin 1979,171.)

Second, sociologists aim to understand the historical developments of scientific episodes, not to evaluate the validity of claims made by participants. Critics do not often recognize how consistently the sociologists of scientific knowledge pursue the program of a value-free social science. In the case of the study of science this means that sociologists fully abstain from evaluating the participants' positions. Evaluative concepts such as truth, rationality, successful, or progressive, are treated as actor's categories. Instead of using normatively loaded concepts like bias, epistemic or convincing, when describing of scientific episodes, a sociologist makes them a topic of analysis. While scientists, and philosophers, regard themselves as participants in a discussion, a sociologist of scientific knowledge takes the position of an outside observer who has no stake in the issue.

Instead of giving her own meaning to the concepts of science and epistemic credibility, she tries to follow how objects of her study themselves understand and demarcate these things, and to determine the consequences of this boundary work. The point is not to maintain, move, or change the borders between good and bad science or between science and non-science, but to observe how these boundaries are created, interpreted and sustained. And indeed, this is one of the most interesting aspects of the SSK. (Gilbert & Mulkay 1984; Gieryn 1999.)

The purely descriptive approach of the SSK has drawbacks. The problem is that the participants in the dispute under study, as well as the outside commentators, often tend to interpret sociological descriptions to be normatively loaded. When a sociologist of scientific knowledge shows that the production of a given piece of scientific knowledge is a social process involving various social interests, outsiders might interpret these claims to be a form of criticism. Especially those supporting the marginal position tend to hijack the sociologist to their side (Lynch 1993, 77-80). And indeed, the descriptions a sociologist of scientific knowledge provides *can* be normatively relevant. The motives of the agents, research practices and standards are what the evaluation of the quality and relevance of research is about. They matter to those who use the knowledge claims and to people who are affected by them. Epistemic evaluation is much like ethical evaluation: it is difficult or impossible to completely step outside of it. Although a sociologist of scientific knowledge wants to remain as an outside observer, she quite often has also some other hats to wear. As a user of knowledge, a tax-payer, a social activist, or as a citizen affected by expert decisions, she cannot maintain the stance of an impartial observer. This is the source of the inherent ambiguity in her position. Descriptive claims made from the observer's perspective can be read as normative comments from the user's perspective. I do not claim that this observation makes the SSK version of value freedom impossible; rather, it calls for more care when publicly presenting sociological studies. What is needed is a clear distinction between the perspectives of an observer of knowledge production and those of a user of the knowledge produced.

4. How Not to Criticize Interest Explanations

I have argued that there is no special ‘interest theory’ in the SSK and that the sociologists of scientific knowledge are in principle using the same folk psychological notion of ‘interest’ as the scientists they study and as the other people in science studies. The only difference is that their notion is not normatively laden or restricted to the non-epistemic goals. In the light of this discussion, it is good to revisit some of the standard criticisms of the interest explanations in the SSK.

There are two kinds of criticisms that stem from two different sources. First, there are the philosophical critics of the SSK who have claimed that the supporters of interest explanations oppose explanation of action in terms of reasons (Laudan 1981, 188-193; Brown 1989, 24-30; Roth 1996, 48-54). Second, there are the fellow sociologists of science who have claimed that interest explanations reify social action and relations in undesirable ways. The users of interest explanations have been accused of taking society for granted (Latour 1987; Pickering 1995, 151-152), treating interests as unmoved movers (Woolgar 1981, 372; Pickering 1995, 63-65) and taking social agents as ‘interest dopes’ (Woolgar 1981, 375).

Let us start with the claimed conflict between interests and reasons. The Strong Program denies the causal and the explanatory relevance of *our* evaluations of the historical agents’ rationality. However, there is no reason to conclude that the agents’ reasons cannot affect their own behavior. In fact, interest explanations *presuppose* that the agents’ practical reasoning causally influences their behavior. There is a clear difference between what an agent believes to be ‘good reasons’, and what an outside observer takes to be ‘good reasons’. And only the former are *causally* relevant to the particular historical episode. This is the true content of the sociological denial of the explanatory relevance of philosophical theories of scientific rationality, and I think that every self-respecting historian of science would agree.

The sociologists of scientific knowledge also want to emphasize that we can try to explain *why* agents accept precisely those reasons as *good* reasons. One should sometimes try to transform one's *explanans* into an *explanandum*. In fact, the sociologists of scientific knowledge think that this is a mark of a serious scientific attitude. After all, the immediate explanation might not be the most interesting. I conclude that this criticism is misguided and should not be regarded as troublesome for the Strong Program. Interests are not mysterious social factors working behind agents' backs and overriding their reasoning processes. Rather, they are things that figure in their practical reasoning.

What about the sociological critiques? Are interests unchanging, unexplainable or are interest explanations treating agents as dopes? I cannot evaluate these claims as criticism of the individual uses of interest concepts here. However, as general claims about interest explanations, these accusations have to be rejected as misguided. There is nothing inherent in the idea of interest explanation that would force one to commit to any of these fallacies. In fact, the interest perspective stems from exactly the same view of scientific work as the goal-directed, situated action that critics are offering as an alternative (Shapin 1988, 544). Similarly, it is difficult to see why an explanation in terms of interests should in any way make it impossible to allow that interests can change in time or that new interests can arise. To the contrary, because interests are related to an agent's action situation and goals, it is difficult to see how this perspective could even deny the possibility of interests changing. It does not follow from the fact that one cannot explain everything at once that certain factors are given the status of unexplained explainers. The interest explanation takes the agent's action situation and her goals as to be given in a particular explanation, but this does not preclude the possibility that these might change or that they might in turn be objects of explanation. Every explanation has to take something as a given, and not even a social constructivist can escape this basic fact about explanation.

What about the claim that the agents are being treated as 'interest-dopes'? This characterization gives a very misleading picture. The interest explanation draws from the fact that an agent's action

situation narrows the avenues of action that are available to her and that the agent knows this. This explanation is not committed to the idea that an agent cannot change her goals or to the denial of the important fact that the formation of an agent's goals is often a social process. Both of these issues can and should be taken into account and analyzed within the interest perspective.

The belief that there is something such as 'the interest theory' is a myth. Like everybody else, the advocates of the Strong Program are applying ordinary common sense when attributing interest. What separates the Strong Program from other schools in science studies is the fact that they have raised the concept of interest to the status of an emblem. This observation makes the seemingly confused nature of the critiques of the interest explanations (Woolgar 1981; Latour 1987) more understandable. They were not so much directed against the concept of interest itself, as against the vision of the goals and methods of sociology of scientific knowledge that it was taken to symbolize.

It is clear that the Strong Program would be much more interesting if it contained a substantial theory about the interests that affect the production of scientific knowledge. One could say that it has partially failed as a research program since it has not been able to fulfill its initial promise. But in defense of the Strong Program, I wish to point out a couple of important reasons for their openness with respect to the kinds of interests that have, or could have, influenced the development of science. First, according to their naturalistic position, it is a matter of historical and sociological investigation to find out what kinds of interests are at work in a certain historical episode (Barnes, Bloor & Henry 1996). The question is empirical and should not be prejudged by philosophical or theoretical arguments. This makes it important to keep the issue open. Secondly, which interests are relevant depends crucially on the intended *explanandum* the sociologist has in mind. (See Ylikoski 2001, 123-130) It would be foolish to limit the list of possible explanatory factors before one has chosen which aspect of the scientific episode one wishes to explain. These two considerations show that the programmatic nature of the theoretical discussion is not a simple failure. There is some justification for the abstractness of the formula 'goals and interests'.

These considerations might lead one to conclude that both instrumentalism and the interest perspective advocated by the Strong Program are rather trivial claims. Even the cognitive goals discussed in the traditional philosophy of science seem to fit the bill, since there are no limitations imposed on the kinds of interests that are allowed to figure in the interest explanations. Furthermore, nobody has denied the goal-directed nature of scientific activities. I take this to be a valid point, at least to a certain degree. It seems that most of the critical discussion of interest explanations in science is based on misunderstandings. Once the misunderstandings are resolved, there is not much to disagree with. However, instrumentalism includes two important ingredients that are missed by the triviality accusation.

The first ingredient is a heuristic for empirical analysis. Instead of simply assuming that all prominent scientists try to achieve the cognitive goals that an analyst believes to be the proper goals of science, the interests perspective suggests that the analyst should ask who the principal agents in scientific developments are and what their characteristic goals and interests are. Although all scientists might have an abstract ‘interest in prediction and control’ (or an equally abstract goal of truth), instrumentalism advises us to look for the differences in the agents’ interpretation of this general goal. It might be that although at the verbal level everybody subscribes to the same ‘official’ goals, yet under the surface, one finds a great deal of variance in the accounts of how these abstract goals are translated into concrete actions. It probably also turns out that the scientists have a number of non-official interests that also influence the direction of their research and the interpretation of their results.

The second important ingredient is the idea that it is both legitimate and important to ask *why*-questions about the cognitive goals of scientists. The SSK does not differ from the traditional history and philosophy of science by replacing cognitive goals by social interests as the central explanatory factors. Instead, the novelty of the SSK is in its insistence that it is also important to explain the cognitive goals that scientists have. Traditionally, students of science have not been

eager to ask why scientists have the scientific goals they have. The SSK emphasizes the importance of this *explanandum*.

These points are now much more widely accepted than in the 1970's when the Strong Program was launched. There is therefore no reason to consider the interest explanations in SSK to be especially radical challenges to our understanding of science. However, this does not mean that the interest perspective has completely lost its relevance to the debates in the social studies of science. As I have argued elsewhere (Ylikoski 2003), the recent 'practice turn' in the social studies of science has been based on the misguided assumption that concepts like 'form of life' and 'practice' are replacements for interests in the analysis of scientific activities. The fact is that 'goals and interests' are still relevant in understanding practices. However, they are just one factor to be considered, not a magic bullet for understanding science.

Acknowledgment

I thank Pekka Mäkelä, Susanna Snell and two anonymous referees for their useful comments.

References

- Barnes, Barry 1977. *Interests and the Growth of Knowledge*. London: Routledge & Kegan Paul.
- Barnes, Barry 1985. *About Science*. Oxford: Basil Blackwell.
- Barnes, Barry, Bloor, David & Henry, John 1996. *Scientific Knowledge. A Sociological Analysis*. London: Athlone Press.
- Barnes, Barry & MacKenzie, Donald 1979. On the Role of Interest in Scientific Change. *Sociological Review Monograph* 27: 49-66.
- Biagioli, Mario 1993. *Galileo Courtier. The Practice of Science in the Culture of Absolutism*. Chicago: The University of Chicago Press.
- Bourdieu, Pierre 1975. The Specificity of the Scientific Field and the Social Conditions of the Progress of Reason. *Social Science Information* 14: 19-47.
- Brown, James Robert 1989. *The Rational and The Social*. London: Routledge.
- Callon, Michel and Law, John 1982. On Interests and their Transformation: Enrolment and Counter-Enrolment. *Social Studies of Science* 12: 615-625.
- Fuchs, Stephan 1992. *The Professional Quest for Truth. A Social Theory of Science and Knowledge*. New York: SUNY.
- Giere, Ronald N. 1988. *Explaining Science. A Cognitive Approach*. Chicago: The University of Chicago Press.
- Gieryn, Thomas P. 1999. *Cultural Boundaries of Science. Credibility on the line*. Chicago: The University of Chicago Press.
- Gilbert, Nigel & Mulkay, Michael 1984. *Opening Pandora's Box. A sociological analysis of scientists' discourse*. Cambridge: Cambridge University Press.
- Habermas, Jürgen 1972. *Knowledge and Human Interests*. London: Heinemann.
- Hirschman, Albert O. 1977. *The Passions and the Interests. Political Arguments for Capitalism*

- before Its Triumph*. Princeton: Princeton University Press.
- Hull, David L. 1988. *Science as a Process. An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: The University of Chicago Press.
- Kim, Kyung-Man 1994. *Explaining Scientific Consensus. The Case of Mendelian Genetics*. New York: The Guilford Press.
- Latour, Bruno 1987. *Science in Action*. Milton Keynes: Open University Press.
- Laudan, Larry 1981. The Pseudo-Science of Science?. *Philosophy of the Social Sciences* 11: 173-198.
- LeGrand, H. E. 1988. *Drifting continents and shifting theories*. Cambridge: Cambridge University Press.
- Lenoir, Timothy 1997. *Instituting Science. The Cultural Production of Scientific Disciplines*. Stanford: Stanford University Press.
- Lynch, Michael 1993. *Scientific practice and ordinary action. Ethnomethodology and social studies of science*. Cambridge: Cambridge University Press.
- MacKenzie, Donald 1981. *Statistics in Britain 1865-1930. The Social Construction of Scientific Knowledge*. Edinburgh: Edinburgh University Press.
- MacKenzie, Donald 1990. *Inventing Accuracy. A Historical Sociology of Nuclear Missile Guidance*. Cambridge: MIT Press.
- Myers, Milton L. 1983. *The Soul of Modern Economic Man. Ideas of Self-Interest Thomas Hobbes to Adam Smith*. The Chicago: University of Chicago Press.
- Pickering, Andrew 1995. *The Mangle of Practice. Time, Agency & Science*. Chicago: The University of Chicago Press.
- Resnik, David 1998. Conflicts of Interest in Science. *Perspectives on Science* 6: 381-408.
- Roth, Paul A. 1996. Will the Real Scientists Please Stand Up? Dead Ends and Live Issues in the Explanation of Scientific Knowledge. *Studies in the History and Philosophy of Science* 27: 43-68.
- Rudwick, Martin 1985. *The Great Devonian Controversy*. Chicago: The University of Chicago Press.
- Sapp, Jan 1987. *Beyond the Gene. Cytoplasmic Inheritance and the Struggle for Authority in Genetics*. Oxford: Oxford University Press.
- Shapin, Steven 1979. The Politics of Observation: Cerebral Anatomy and Social Interests in the Edinburgh Phrenology Disputes. *Sociological Review Monograph* 27: 139-178.
- Shapin, Steven 1988. Following Scientists Around. *Social Studies of Science* 18: 533-550.
- Shapin, Steven 1992. Discipline and Bounding: The History and Sociology of Science as Seen Through the Externalism-Internalism Debate. *History of Science* 30: 333-369.
- Shapin, Steven & Schaffer, Simon 1985. *Leviathan and the Air-Pump. Hobbes, Boyle, and the Experimental Life*. Princeton: Princeton University Press.
- Segerstråle, Ullica 2000. *Defenders of the Truth*. Oxford: Oxford University Press.
- Tammi, Timo 1999. On Experimental Discourse in Economics. *Philosophy of the Social Sciences* 29: 62-89.
- Whitley, Richard 1984. *The Intellectual and Social Organization of the Sciences*. Oxford: Clarendon Press.
- Woolgar, Steve 1981. Interests and Explanation in the Social Study of Science. *Social Studies of Science* 11: 365-94.
- Yearley, Steven 1982. The Relationship Between Epistemological and Sociological Cognitive Interests: Some Ambiguities Underlying the Use of Interest Theory in the Study of Scientific Knowledge. *Studies in the History and Philosophy of Science* 13: 353-388.
- Ylikoski, Petri 2001. *Understanding Interests and Causal Explanation*. PhD. Thesis. University of Helsinki. Available at <http://ethesis.helsinki.fi/julkaisut/val/kayta/vk/ylikoski/>
- Ylikoski, Petri 2003. Explaining Practices. *Protosociology* 18: 316-330.