5

On Interviewing "Good" and "Bad" Experts

Jochen Gläser and Grit Laudel

5.1 The "quality" of interviewees matters

The success of interview-based investigations considerably depends on the "quality" of the interviewees, that is on the extent to which they meet our expectations in the interview situation. We expect interviewees to understand which information we need, to provide this information in extensive, complete and detailed responses, and to adjust their communication to our steering of the conversation. We also hope to meet respondents which reflect on their own social situation and who are able to provide information about their perceptions, social relations and motives.

These aspects of "quality" are particularly important in interviews that are intended to "mine" a respondent's special knowledge about a social situation or a social process. We consider all interviews that have this function to be "expert interviews." Thus, we define "experts" as people who possess special knowledge of a social phenomenon which the interviewer is interested in, and expert interviews as a specific method for collecting data about this social phenomenon. This conceptualization of the "expert interview" is based on the expert role of the interviewee in the interview. Thus understood, expert interviews are a distinct method that is applied in investigations of a specific type, namely investigations that reconstruct social situations or processes and use interviewees as a source of information (Gläser and Laudel, 2009, pp. 11–14). This understanding of "expert interviews" has been introduced to the German methodological discussion by Hopf (1993). In the Anglo-American literature, the distinction between the roles of interviewees as sources of information respectively objects of study is reflected by the concepts of "informants" and "respondents" (for an early reference to this distinction see Zelditch, 1962).

Another approach to "expert interviews" is their conceptualization as interviews with people who have an expert role in the investigated social setting. In that perspective, which dominates the wider literature as well as the contributions to this volume, experts are people who are set apart from other
actors in the social setting under investigation by their specific knowledge and skills. Such superior knowledge is usually produced by designated processes of learning and training (for example vocational training). Members of professions such as physicians, lawyers, or architects are the best-known examples of "trained" experts. However, expert roles in social settings are not limited to the professions.

This understanding of experts ties the concept "expert interview" to a specific kind of respondents or informants but does not limit the forms or functions of interviews with them. It also focuses the methodological interest on the particularities of an interview that result from the expert status of the interviewee in the investigated field. Thus understood, the expert interview is not a specific method of data collection but includes all forms of qualitative interviews that are conducted with experts. In our opinion, this is the reason why there is no specific methodological discussion of the expert interview as a method of data collection, as has been noted by Bogner and Menz (Bogner and Menz, 2005, pp. 11–16).

Since the roles of "interviewees as experts" and "experts as interviewees" are constructed in different social situations (the interview respectively the investigated social settings), the two definitions of expert interviews are not disjunct (Figure 5.1). In many social science investigations only one of the roles is relevant. The left (white) area in Figure 5.1 represents investigations in which interviews are aimed at the reconstruction of social processes and situations but are conducted with people who don't have an expert role in the field. For example, interviews about health care could be conducted not only with physicians but also with patients. The latter don't have an acknowledged expert role in health care but possess special knowledge about the way health organizations handle patients and thus would be interviewed as experts for this aspect of the investigation. The right (black) area stands for interviews with people who have an expert role in the investigated social field but who are not interviewed as sources of information about social processes in which they participate. An example of this would be a study of biographical self-representations of physicians, where the reflection of social processes by physicians rather than the processes themselves are the subject matter of the investigation. In the middle (grey area) we find the area of overlap between the two definitions. The overlap refers to interviews with "experts in the field" who serve as sources of information about a part of the social reality.

For this specific group of interviews an additional aspect of the "quality" of interviewees becomes important. If we use "experts in the field" as sources of information, our interviews depend not only on their performance in their role as interviewees but also on their "quality" in their expert role in the field. We must ask ourselves whether we receive the same information about medical treatments from good and bad physicians, the same information about baking from good and bad bakers, and the same information about competitive sports from world champions and average athletes.

This problem has not yet explicitly been dealt with in the literature even though it might be implicit to discussions about interview responses. We address it in this chapter by using examples from our own empirical investigations, in which we must compare and synthesize statements of researchers. In doing so, we face the above-described quality problem: Do "good" researchers describe their situations in other ways than "bad" ones? What does it mean that certain issues are described in the same way by good and by bad researchers? How can we find out how "good" our informants are in their researcher roles?

5.2 "Good" and "bad" researchers as information sources of science studies

There can be little doubt that researchers are "experts in the field." It also is no secret that the abilities of researchers differ and that the quality differences between very good ("excellent") and medium or even bad researchers are huge. Nevertheless, science studies hardly ever took the abilities of the investigated researchers into account. The abilities of researchers who act as informants in qualitative investigations have not been a topic of methodological reflection at all. This may be due to the fact that outsiders are hardly able to judge the abilities of a researcher. Indeed, there are only very few researchers who are regarded as outstanding and are well-known outside science, such as Nobel price winners and - to a lesser extent - members
of the learned academies. Below this threshold of visibility the quality of researchers remains hidden in the implicit judgements of their scientific communities.

Two kinds of empirical science studies can be distinguished according to their handling of the quality problem. Most studies ignore the problem of quality differences between respondents. The first interview-based investigations are classical examples of this neglect (for example Hagstrom, 1965, Crane, 1972). The subsequent ethnographic studies also ignored the “quality” of the observed researchers and possible influences of this aspect on the outcomes of observations (Knoer-Cetina, 1984, Lynch, 1985, Latour and Woolgar, 1986 [1979]). This is surprising because we cannot a priori assume that the observed processes of the social construction of knowledge is performed the same way by “good” and “bad” researchers. The laboratory studies are implicitly based on this very assumption because they do not include the quality as intervening factor – in spite of the obvious importance of researchers’ abilities to the observed scientists. The quality discourse among the observed scientists was reported (Latour and Woolgar, 1986 [1979], pp. 163–65) but the quality problem neither influenced the empirical strategy of laboratory studies nor their interpretation. Since the constructivist tradition – which constitutes the mainstream of the sociology of science since the 1980s – still disregards the quality of the makers of scientific knowledge, we still don’t know for sure whether good researchers construct knowledge differently from bad ones.

A second group of studies chose research performance as its topic and made the performance levels (productivity, creativity) of researchers the dependent variable of their investigations. The aim of these studies was to identify influences on the productivity and the quality of scientific work. While investigating the influence of organizational research conditions on the productivity of scientists, Pelz and Andrews included the quality – measured by the number of publications – as a dependent variable (Pelz and Andrews, 1966). These kinds of studies were taken up later in the context of investigating scientific excellence and scientific creativity (Jackson and Budlton, 1987, Zuckerman and Cole, 1994). In her study of Nobel Prize winners, Zuckerman did not encounter the problem of quality differences because the study was limited to a homogeneous group of a visible elite (Zuckerman, 1977). Thus, she did not address quality differences in her methodological reflections but instead wrote about the specifics of interviews with members of the elite (Zuckerman, 1972).

With the development of the Science Citation Index it became possible to use a more sophisticated measure of quality, namely the frequency with which publications are cited by other scientists. This measure was first introduced by Cole and Cole who started to treat quality not only as a dependent but also as an independent or intervening variable (Cole and Cole, 1967, Cole, 1970, Cole and Cole, 1972). Their use of citations created a new research tradition within bibliometrics for which the measurement of the quality of researchers, organizations and national science systems is an important topic. Shortly thereafter, research policy became interested and began to generate an ever-increasing demand for performance measurements. Today these measurements also utilise other than bibliometric indicators, particularly the amount of external research funding. All these indicator measures are somewhat problematic, only work in some research fields, and are of questionable validity at the level that is of interest here, namely that of the individual researcher (Laudel, 2005, Gläser and Laudel, 2007a).

This short overview demonstrates that science studies have either ignored the varying performance levels of researchers or turned them into the subject matter of their investigations. The methodological question about different responses from “good” and “bad” researchers and the consequences of such differences for studies that are based on interviews with researchers has not yet been asked.

5.3 The dependence of data on the performance levels of researchers

5.3.1 Different phenomena, different perceptions

In our own research we are interested in the influence of institutional and institutionally produced conditions of research on the construction of knowledge production. We use qualitative interviews with researchers and research managers as the main method of data collection. We ask researchers about their research projects (the problems they are working on, the methods and objects they use, their collaborations, and so on) and about the specific resource requirements of these projects. Another set of questions focuses on the conditions of research, particularly the available time for research, access to resources, and possible influence of the organizational environment on the content of research. From the responses to these questions we reconstruct the changes in the content of research that result from the adaptation of topics, objects, methods, and collaborations to the conditions of work.

We had to learn that in our investigations talking to “good” or “bad” researchers makes a difference. The “quality” of the interviewees influences our investigations at two levels. Firstly, the situations of “good” and “bad” researchers may differ, for example with regard to the availability of external funding, the workload, or the status of the interviewees in their organization. In this case, the researchers experience different phenomena. Secondly, it is possible that “good” and “bad” researchers provide different descriptions of the experienced phenomena because their perceptions or communication behaviour differ. If we combine these two variations we can construct four variants of communication in the interview (Table 5.1).
Table 5.1  Quality-dependent communication situations in interviews with researchers

<table>
<thead>
<tr>
<th>Phenomena</th>
<th>Same</th>
<th>Different</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>(False) implicit premise of science studies</td>
<td>Appears as contradiction in the responses. Main problem: identifying the &quot;real&quot; situation</td>
</tr>
<tr>
<td>Different</td>
<td>Not visible in the responses. Main problem: Recognising that a difference exists</td>
<td>Most common situation. Main problem: Recognising how the data are influenced by quality differences</td>
</tr>
</tbody>
</table>

The left upper cell of the table describes the implicit premise of science studies, namely that all responses are equal in both dimensions. Of course, nobody has ever said that this is actually the case. However, disregarding the quality problem in interview-based and ethnographic studies means that data collection and analysis is in fact based on this presumption.

The common situation in science studies is the direct opposite of this presumption. It is situated in the right lower cell of the table, indicating that the "good" and the "bad" researchers experience different things and provide descriptions of their experiences that are influenced by their "quality." For example, a historian explained during the interview that he can freely choose his topics because the publishers take everything that he writes. In contrast, another historian described that his book project has been rejected by several publishers, and therefore he will now change the topic of his research. The publishers neither requested reviews about the book project from peers nor did they tell him the reasons for the rejections. He assumed that the topic was regarded as not profitable. In the course of the ethnographic observation both historians were interviewed. We found a clear quality difference between the two historians, which let us assume that the insufficient quality of the book project was the reason for the repeated rejections and that the reference to the topic as reason was merely a rationalization of the respondent. However, the ethnographic observation also showed that publishers indeed reject books for thematic reasons. Given that the two historians worked in two entirely different areas, a mix of different experiences and quality differences occurred that was difficult to untangle.

In the "standard case," the complexity of an analysis is dramatically increased because two dimensions must be added in the interpretation of interview statements. We avoid this case in our following demonstration of the problem because its discussion would require a large amount of background knowledge that cannot be provided in a book chapter. Instead, we use the two "pure" cases that only vary in one of the two dimensions. In our experience the pure cases are rare – it does not often happen that "good" and "bad" researchers give identical descriptions of different phenomena or different descriptions of the same phenomena. However, in our current investigation we conducted enough interviews to obtain these kinds of responses. We use them in order to demonstrate how difficult the standard situation is. It is the right lower cell of the table that describes the real challenge because in most situations we are confronted by different descriptions that are likely to point to different phenomena. It is up to us to determine how the performance levels of interviewees influences their descriptions.

5.3.2 Identical descriptions of different phenomena

Our interviews contained several questions that targeted adaptations of the research content to the resource and institutional conditions of research. We intentionally phrased these questions in an indirect and open way in order to avoid an orientating influence on the interviewee. One of the questions was: "Are there research topics that you would like to work on but can't?" The answers to this question varied considerably. The most frequent answer was yes, there are indeed research problems that would be nice to work on but there is not sufficient time. Other interviewees mentioned their resource base as limiting factor. A third group did not feel any restrictions. Among those were two historians who worked at the same university. One of them is a professor who has published numerous books that are read and cited beyond Australia. He regularly acquires external research funding and is a member of the Australian Academy of the Humanities. The other historian obtained his doctoral degree six years ago, but has neither published his PhD thesis nor any other substantial research result. He has never acquired external research funding and was unable to describe research projects of his own even after being repeatedly asked in different phases of the interview. Who is who?

**Interviewer**: And are there research topics that you are interested in but can't work on?

**Historian 1**: No. I've been very lucky.

**Historian 2**: ...that I'm interested in but I can't work on? No. No, I think there is still ample freedom to pursue that curiosity driven approach. If I was to be seized with a particular idea, the resources are here, all the resources are available [...], because I think with historical research, the demands, the cost issues, are relatively modest. It's really my own time, photocopies, inter-library loans and travel. Put all those together and that's really all you need for historical research. So, for that reason, I think most projects that I would be interested in pursuing, are always
going to have fairly modest resource implications, and that’s definitely no real reason why I couldn’t pursue them.

Both interviewees perceive a situation, which does not limit their research financially, institutionally or in any other way. In both cases the aspiration levels concerning the conditions of research correspond to the actual conditions of research. However, the satisfaction of both interviewees is situated at significantly different levels. One of the interviewees has sufficient time and resources for the projects he wants to conduct (not the least because he can use the external funding to reduce his teaching load), while the other one doesn’t conduct projects and therefore also perceived a correspondence between his interests and his conditions of work. The difference between a situation with sufficient external funding and a situation in which even the recurrent funding is only intermittent and based on internal grant proposals is not visible in the two assessments of the respective situations.

Researchers from other disciplines also gave near-identical answers that resulted from different aspiration levels. One of the two mathematicians and one of the two geologists rarely publish, are hardly cited, don’t acquire external grants and don’t have a clear research programme, while the other two publish a lot, are cited above average, continuously get grants and pursue long-term research programmes.

_Interviewer:_ Now, are there any research topics that you are interested in but can’t work on?

_Mathematician 1:_ Not really, not really at all. I mean I’d like to have, I’m hindered in the financial maths to some extent by not having someone to talk to. Well, there is one other person here that does work in it but he works on it from a different angle so, in some respects I would like to do more in financial maths, but on the other hand I’d like to do more in symmetries because that’s what I really like to do. But, no, I don’t think that there’s an area that I can think of that I really want to do.

_Mathematician 2:_ I don’t think so. No, I feel at the moment I have quite a few research topics and I’m not looking for – I’m not actively looking for more. I mean, often research projects just arise naturally. You just get into them because you’re looking at some problem that leads into something else. So I don’t sort of look to start a project from scratch. It always comes from another source. No, I just follow what happens.

_Geologist 1:_ I guess, at the moment it’s sufficient to keep me going at this stage. I think I’ve got sufficient projects to keep me occupied at this stage. So I’m not really thinking about other opportunities.

_Geologist 2:_ No, not really, I mean everything I am interested in...I can maintain this balance between the applied and pure aspects enough to keep me going now.

We found yet another example of identical descriptions of different phenomena in the self-evaluations of research performance that was sometimes given by the interviewees. The questions about the performance expectations of universities and evaluations of research performance inspired two political scientists at the same university to reflect on how they meet the expectations of their university. One of the two publishes irregularly (with interruptions of up to four years). His only journal publications are book reviews. The other publishes regularly, writes articles in refereed journals, and is internationally visible.

_Interviewer:_ So, my last question about the evaluations would be: What are the consequences of these evaluations for you?

_Political Scientist 1:_ In terms of, obviously, promotion. It’s whether you get promoted or not. In terms of professional development – _I mean, I am not highly productive but I produce on a reasonably regular basis._ I’ve never been in a situation where someone said, you haven’t produced anything and you’re going to. I can always say I’m in the middle of a book now and it’s going to come out next year or the year after. So I’ve never been in a position where anyone’s questioned me so I don’t know on what basis they would.

_Interviewer:_ Are there also specific special expectations in terms of the quality of your research from the University or School?

_Political Scientist 2:_ Well, no there is always some talk I guess again, [...] As far as I know there is nothing on that type of functions. And it is clearly that for most people in all universities never minc in [our university], where it’s the number of publications that seems to matter more than quality. And I don’t see any staff pressure counteracting that. _I don’t publish very much._ Even though I keep count of my rates, which is now slowing down because of too much teaching. I don’t think I feel affected by any pressures either way. I am aware of the pressures, but only a small number of us can do it.

The two passages in italics highlight the problem: without additional data about the publication activities of respondents we would consider both interviewees as not being particularly productive and as equally productive. We would even consider the first ("I produce on a reasonably regular basis") as the more productive and the second as less productive. But it is exactly the other way around. Now let us imagine the many descriptions of this kind that we cannot check independently of the interview...

5.3.3 Different descriptions of identical phenomena

We encountered the other of our pure cases in description of evaluations, that is of situations in which the university assessed the quality of its
academics and based decisions on these assessments. Among these evaluations, those that were conducted in the contexts of internal research funding and promotions had the strongest influence on the situation of researchers. The evaluations for promotions are based on a university-wide standard procedure that is only slightly modified in order to accommodate differences between disciplines. The following example includes two historians from the same university who faced the same procedure and evaluation criteria when it came to promotions. From the quotations it is clear which interviewee is "good" and which is "bad." It is far less clear that they both describe the same procedure.

Interviewee: And is this the only occasion when your research is evaluated? When you apply for internal grants?

Historian 1: No, it's evaluated if I was to submit an application for study leave, which I have done, and to submit an application for promotion. And I prepared an application under the previous Dean back in 2000 and the then Dean made it quite clear even before I submitted the application that she wasn't going to support it because the first thing she said to me, I had a meeting with her, was my research profile is not very good.

Interviewer: And since then you didn't try again to...

Historian 1: I revised my application and again got as far as having a meeting with the Dean last year but it was indicated that I would be... And my Head of School has indicated to me that I can be sure of being successful in my application for promotion, this is only to Lecturer C, once I submit the book manuscript. So in other words he's telling me that I will only be successful in promotion from a Lecturer B to a Lecturer C if I publish the book, which to my mind is... We have people in the faculty and elsewhere who've been promoted to Associate Professor and we've had people who've been appointed as Professor who don't even have a book and yet I've been told as a condition of my application for promotion being successful that I have to have submitted a book manuscript which seems to me unfair.

Interviewer: And, how does the university evaluate your research?

Historian 2: Okay. Well, promotions would be the major mechanism. [...] Yes, so on the individual level it's promotion and career development interviews.

Interviewer: What are the consequences of these evaluations for you?

Historian 2: Well, in my case, because I'm quite productive, it's usually good in the sense of getting promoted. But at present there haven't been negative consequences. But if, say, within the Faculty, people are not being very productive, then they're given mentoring, and I've done that as well. You engage with people and you advise them on what they could do to get together a great application or what kinds of publication strategies they need to have.

Both descriptions refer to the same promotion procedure. Academics write an application in which they describe their achievements in teaching, research, and academic administration. They discuss this application with their supervisors (first with the head of school and thereafter with the dean of their faculty), who advise them on the content of the application and on their chances of success. When the application is moving up the hierarchy, each supervisor adds their comments. The final decision is made by a university committee. If the supervisors believe the application would be unsuccessful they advise academics to withdraw it. Such a withdrawal is in the interest of both the applicant and their supervisor. Academics whose application is rejected must wait several years before they can apply again, while postponing an application that has little chance of success enables another application with better chances in the following year. A weak application that goes ahead also creates a dilemma for the supervisors. If they support it, they lose their credibility at the higher levels of the hierarchy. If they don't support it they lose their credibility with the applicant.

The application procedure is strictly regulated. The rules are made publicly available and are strictly followed. Both interviewees we quoted above wrote their applications and had consultations in which they received advice. However, their perceptions of the process are quite different. One interviewee mentions his promotions only casually and gives the impression that they were not problematic. For him, the stronger and more present experience is that of his involvement as a mentor, a role, which has been introduced to support weak applicants. The other interviewee applied with achievements that were considered as insufficient by two deans and the head of school. The supervisors advised against an application. The applicant regarded this as an unfair interpretation of performance criteria, a rather unlikely interpretation given the strict rules and the consistent judgment by the two deans and the head of school. However, it is interesting to note that something that is interpreted by the "bad" applicant as an unfair intervention that prevented the promotion is of no personal concern to the "good" applicant, who hardly even mentions it. The different performance levels of the two interviewees have a massive impact on their perceptions and descriptions of what is essentially the same phenomenon, namely the promotion procedure. This would not be a problem if only their perceptions would matter for the investigation. However, we also need to know the promotion procedure. While in this particular case the "true" procedure could be established by triangulation from many different sources, it is again easy to imagine situations where this would be impossible, and quality-dependent descriptions were the only information about a phenomenon.
5.3.4 Implications for the investigation

In our comments on the quotations from interviews we have already hinted at their different informational yield for our project. Now we would like to systematize these consequences of the quality problem. The differences in the performance levels of our respondents shaped their communication about two aspects of their work that were central to our investigation. The first aspect is a respondent’s content of work. When we ask for an explanation of the respondent’s research problem, reasons for the selection of specific research objects and methods, or other information about the content of research, “bad” researchers provide information that is different from that provided by “good” researchers. This information about the content of work is crucial for our investigation because the content of work shapes the resource demand. “Good” researchers formulate more challenging research problems, and their assessment of strategic aspects of their research such as the solvability of problems or the applicability of methods to their problems differs from assessments of the same aspects by “bad” researchers. These strategic choices and assessments produce the specific resource requirements of the respondent’s research, which constituted one of the central variables of our investigation. This is why the results of our study partly depend on what the biochemists tell us about their biochemistry, geologists about their geology and political scientists about their political science.

In addition to the specific information about the respondents’ own research the interview also yields general information about the characteristic epistemic practices of their research field. Some characteristics of fields (for example the typical duration of research processes, specific uncertainties or heterogeneity of the used knowledge claims) are important intervening variables in our investigation because they explain the varying implications of uniform resource allocation mechanisms in different research fields. This is why we must extract such field-specific factors from the statements of all respondents belonging to the same field. A description of field-specific practices must be synthesized from statements by researchers of varying performance levels.

This “learning from the field,” that is the use of experts as a source of information about characteristics of the content of work, occurs in many science studies projects and probably in many other studies that are based on expert interviews. It is always a means to an end because the knowledge about work processes of experts is needed to answer sociological questions. This is why a second aspect of our respondents’ work – the conditions of work under which the research was conducted – was more important to us than the content of research, and why most of our examples refer to these conditions. The performance levels of respondents strongly affected the data about conditions of work. Our examples illustrated that

- the statement “my research is not constrained by limitations to my access to resources” may reflect two completely different situations, namely one of a researcher whose aspiration level is so low that it can be met by any resource base, or one of a researcher who is so “good” that they receive all the resources they ask for; and that
- the perception of organizational routines largely depends on the performance levels of respondents, that is that largely identical routines may be differently perceived by “good” and “bad” researchers.

For these reasons we could not simply extract general, university-specific or faculty-specific conditions or routines but also had to include the performance levels of interviewees in our interpretation. We could not avoid this problem by relying on documents describing the formal bureaucratic routines because the actual use of these routines in everyday university life (which mediates the influence of institutions on the content of research) inevitably deviates from the formal prescriptions in documents.

If we tentatively generalize our experiences we find that quality differences between experts interfere with two essential tasks of expert interviews:

1. If we need to learn something about the content of an interviewee’s work, interviews with “good” experts provide information about this work that is different from that provided by “bad” experts.
2. If we need to learn something about the interviewee’s conditions of work we obtain quality-affected information for two reasons. Firstly “good” experts have higher aspiration levels and higher standards against which conditions of work are assessed. Secondly, interventions by the environment are dealt with in different ways by respondents depending on their performance levels, and are thus different for “good” and “bad” experts.

Investigations that rely on at least one of these kinds of information have to face the quality problem.

5.4 Quality differences as a methodological problem of expert interviews and their analysis

5.4.1 How do we assess experts?

Whenever the performance levels of the expert affects their responses, we must know these performance levels and take them into account when interpreting statements. However, a quality assessment of experts seems to be an impossible task. After all, we interview experts for the very reason of accessing their special knowledge, that is knowledge we don’t possess. How could we possibly assess how much of this knowledge they have and how well they are able to use it? Apparently there are only few exceptions to this dilemma. Competitive athletes, for example, are subject to a public
performance assessment that is comprehensible outside their group because they compete against each other.

On first glance our own task – to assess the quality of researchers – puts us in a quite comfortable position, too. The demands of science policy and management have pushed the development of evaluation methods, which by now have widely diffused throughout public science systems. Unfortunately, this does not make our task much easier. Science studies agree that ultimately only researchers' peers are able to judge the quality of their work, and that even this peer review may get it wrong. "Objective" indicators that can be applied by outside observers because they abstract from the content of research only offer partial insights into the quality of research, and evaluate past rather than current performance. Therefore, they can only be used as additional information in a peer review. Moreover, we have already pointed out that quantitative indicators cannot be regarded as reliable at the individual level that we are interested in.

The limited opportunities of performance measurement are highlighted when we ask which aspect of quality we actually intend to measure. Is it the current performance of the investigated researchers or their ability to perform? In our investigation we came across two groups of interviewees for which this difference affected assessments. The careers of some researchers were not yet long enough to enable conclusions about extraordinary research abilities although there were some indications of that. The opposite case occurred, too. Researchers who had made outstanding contributions in the past had ceased to do so at the time of the interview because they didn't receive external funding anymore, had moved away from the mainstream of their field, or were close to retirement and didn't want to start new projects. In both cases it is likely that the scientific abilities of the interviewee exceed the level suggested by the current performance. For our purposes, a categorization according to abilities is more important than one according to the current research performance. However, the opportunities to assess abilities are even more limited when current performance is likely to offer a distorted picture.

Given the numerous problems of performance measurement in science, which we have discussed in our own publications, we had to be extremely careful in our categorization of researchers. We used as many indicators as possible, supplemented them with "independent" information from the interviews, and compared the researchers of one discipline to each other.

Prior to the interview we downloaded information about research interests and projects, indicators of peer esteem such as prizes for research or memberships of learned academies, publications, citations and external grant acquisition from internet pages, the Web of Science, and other publication databases. All these indicators are discipline-specific (for example, citation analyses are extremely problematic in the social sciences and the humanities) and can only be interpreted by taking the context into account.

In fields with a high degree of collaboration it is not always easy to ascribe achievements to a specific researcher. Researchers may profit from being a member of a strong group and participate in publications and grant acquisitions without making outstanding contributions themselves. We addressed this problem by checking indicators for all co-authors of highly cited publications and co-applicants for grants.

This information was used for preparing the interviews. The preparation already included a preliminary assessment of performance levels and the selection of an appropriate strategy for handling the quality problem in the interview. We could not rule out that research was a sensitive topic for many of our interviewees whose performance levels were considerably lower than those of their colleagues. This is why we applied a specific strategy when interviewing academics with apparently sparse research records. We didn't start with questions about research, which would have presumed that the interviewee conducts research as everybody else, but with questions about teaching. In this initial conversation, we established the extent to which the interviewee actually conducts research, and steered the interview towards research activities depending on that information.

In our interviews we contributed to the assessment of quality because our questions about the research biography, conditions of research and future research plans touched a trait of a researcher that is closely linked to quality. We obtained clear indications concerning the aspiration level of the interviewed researchers, that is their standards concerning their own work, their collaborators, other colleagues, and their conditions of research. Thus, the descriptions of the research and conditions of work provided in the interview are important indicators of quality themselves. Another indicator we could use is the consistency of the research topics. Do interviewees pursue a consistent research programme, or are they merely responding to external stimuli such as demands by collaborator or supervisors? Does the interviewee independently conduct research at all?

After the interview we refined the information about our indicators and collected additional information from the internet and publication databases. Using the indicators and information from the interviews, we compared the researchers of each discipline with each other and constructed three groups: a high-level group (16 researchers across all disciplines), a medium-level group (79) and a group of academics who hardly conduct any research (23). In order to highlight the borderline cases whose classification was most difficult, we specifically identified those academics in the medium-level group who were close to either the top or the inactive groups. We conducted the categorization together (and largely achieved agreement). The advantage of this collaborative categorization exercise is the necessity to verbalize impressions from the interview (which only one of us had prepared and conducted). These impressions could thus be discussed in the light of the other information that was gathered independently of the interview.
5.4.2 How do we include the “quality” of our experts into the data analysis?

After we got an idea which of our interviewees are “good,” “average” or “inactive” researchers, we coded this categorization in the file names of the interview transcripts, which means that this information appears in the source tag of each interview statement that was used in the analysis.

Always knowing how “good” a source/expert is helps but doesn’t solve the problem. Being well aware of the problem that the interpretation of the data can never be fully described, we will nevertheless try to demonstrate how we used the information about quality in our analysis. We already mentioned the two possible ways of using the quality information in our short overview of the treatment of research quality by science studies. Firstly, knowledge about the quality of the source must be used for the interpretation of the data. Secondly, in our investigation we also need to know whether the adaptive behaviour of “good” researchers differs from that of “bad” ones. Thus, we must also include the performance level of a researcher as a subject of the investigation.

The way in which we take into account the performance levels of respondents in the interpretation of the data can only partly be made explicit. We are sure that we read interviews and interpret data “differently” because we have an opinion about the performance levels of the sources. This is also a danger for the interpretation if a prejudice against a group occurs. We counteracted these influences by conducting the first step of the qualitative content analysis – the extraction of raw data – prior to determining and coding performance levels. In the subsequent analysis, we used a “check list” and determined whether

- information is shaped by the ‘aspiration level’ of the interviewee (in our case: for example if above or below average expectations concerning research performance and conditions of work may have influenced a statement);
- a specific content of work that is due to specific performance levels might have shaped the information (in our case: for example if above or below average research performance manifests itself for example in the number of parallel projects a researcher works on or in the continuity of the respondent’s publication record),
- differences between performance levels are accompanied by different self-images and valuations (in our case: for example whether different research performance is accompanied by varying assessments of tasks in teaching); and
- whether an interviewee’s performance level affects reported causalities (in our case: for example whether unfavourable research conditions are used in a specific way to explain lower research performance).

These questions are particularly important when we find contradictions in the data or when information that is central to the investigation only stems from few sources. An identification of possible influences of performance levels does of course not mean that we discard the data as “unreliable.” We are just in a better position to decide the role we assign a certain piece of information in the construction of our explanation. The extent to which critical aspects of our typologies, generalizations and explanations could be influenced by performance levels of the sources becomes visible.

Information about performance levels of respondents was not only used for the interpretation of data but also in a second, more direct way. In our investigation, the quality of the researchers is also an intervening variable. If researchers adapt the content of their research to institutional conditions then we must find out whether the adaptive behaviour of “good” researchers is different from that of “bad” ones. We achieve this by constructing categories of performance levels and separately investigating the adaptive behaviour for each category. An example of that strategy is Grit Laudel’s comparative study of conditions for external fund acquisition and their impact on the content of research (Laudel, 2006a, 2006b). Studies of hindering conditions of research frequently face the objection that the complaints of researchers are just the “rationalizations of the losers,” that is that “bad” researchers use unfavourable research conditions as an excuse for performing badly rather than admitting their insufficient abilities. The comparative study of external fund acquisition of experimental physicists in Germany and Australia took this argument into account by distinguishing two performance levels (“top” and “other”) and two categories of the scientists’ resource bases (“rich” and “other”). A cross-comparison of the four groups showed that all four fields were populated (Table 5.2). There are the expected concentrations in the lower right cells, which indicate that external fund acquisition indeed is related to the performance level of a researcher. The existence of “top other” and “other rich” scientists shows, however, that quality is neither necessary nor sufficient for a good external fund acquisition. This already puts the “rationalization argument” on doubt. Additionally, an analysis of adaptation strategies

<table>
<thead>
<tr>
<th>Amount of external funding</th>
<th>Germany</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;Rich&quot;</td>
<td>&quot;Other&quot;</td>
</tr>
<tr>
<td>&quot;Top&quot; scientists</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>&quot;Other&quot; scientists</td>
<td>7</td>
<td>23</td>
</tr>
</tbody>
</table>

Source: Based on Laudel, 2006b, p. 382.
showed their occurrence across all performance levels as a reaction to insufficient resource bases. Thus, the “rationalization argument” could be refuted.

The second example is taken from our current investigation that provided the material for this article. During a preliminary analysis of the interviews we discovered that the adaptative behaviour of “good” and “bad” researchers to the conditions of research funding differed. There is a small group of “top” researchers who

- work on a comparatively broad range of research topics and thus could adapt to political demands by selecting politically desired (which means in Australia application-oriented) topics but
- don’t have to surrender to this pressure because they receive the external funding for their favourite projects even if these projects do not correspond to the political orientations of the research landscape (their high performance being the substitute for “therapeutic correctness”).

In the case of these researchers two things came together: a high thematic flexibility and a low pressure to adapt. At the other end of the spectrum are researchers who experience a high pressures: pressure because their current research projects are not successful. However, these researchers are not able to change their research in the required way because their low performance levels also mean that they cannot apply their accumulated knowledge to a new topic.

This observation made us formulate a hypothesis about the relationship between performance levels and adaptative behaviour (Figure 5.2).

Adaptations to institutionalized expectations concerning the content of research occur neither at the top level (because these academics perceive little pressure), nor at the weakest performance level (because these academics are not able to make the necessary changes of direction even under strong pressure). The institutionalized pressure to adapt has its main impact at the medium level of academics who are “not good enough” to protect themselves from pressure through “excellence” but good enough to respond to this pressure with changes to the content of their research.

5.5 Concluding remarks

To what extent can our considerations about the role of the quality of experts be generalized? We believe that the problem of “quality-dependent” data in expert interviews affects many investigations because the defining property of an expert – specific knowledge and abilities – varies between experts. We therefore regard it as necessary to consider possible influences of quality differences between experts at the beginning of an investigation. If data may be affected, appropriate strategies for the selection of experts, the conduct of interviews and the data analysis must be developed.

Depending on the aim of a study, the performance levels of experts can also play a role beyond the “area of overlap” identified in the introduction. We might for example be interested to know whether biographical self-representations of “good” physicians are different from those of “bad” ones. In this case, the quality is explicitly part of the subject matter of the investigation and must be dealt with. Quality differences occur as a “hidden” methodological problem only in studies where “experts in the field” are sources of information about social situations and processes they observed.

Although the techniques of collecting information vary with the research problems and the experts involved, the steps we outlined can presumably be generalized. The main point is to decide early whether the quality of experts could be a problem and whether the strategy of interviewing must be adapted. This includes the decision about using the interviews for collecting data on indicators of quality such as the respondent’s aspiration level concerning their own work and conditions of work. With regard to the data analysis a decision must be made whether the quality of experts could have influenced the information they provided. If this is the case, the strategy of data analysis must be adapted accordingly. Our questions about the aspiration levels, work load, self-image, assessments, and rationalizations seem to be applicable beyond interviews with researchers.

Researchers are an interesting “show-and-tell piece” because quality differences between them are a generally recognized and in part publicly negotiated fact, and because quality-relevant information is publicly available. In other investigations that are based on expert interviews it may be far more difficult or even impossible to get “objective” information (that is
information that is independent from the interviewer’s judgement) about the “quality” of experts. As is the case with all methodological problems in social research, the first necessary step is to emphasize the problem and to accept its relevance for the investigation. It was our aim to initiate this step.

Notes


2. A description of our project and the empirical methods can be found in Gläser and Laudel (2007a).

3. Unfortunately, it is nearly impossible to find out how information on the content of research is influenced by the performance levels of the interviewees because each research process is unique. Differently from bakers, physicians or managers whose work is repetitive to a large extent, researchers aim at producing new knowledge, which makes each research process unique. Consequently, there is no solid base for a comparison of research processes. Even in the case of collaborative research processes whose participants could all be interviewed, the comparison is hindered by the unique disciplinary perspective of each collaborator. PhD students are the only exception to this “problem of uniqueness” because their supervisor is able to assess the research process as well as the researcher.

Further readings


References


Research Methods Series

Interviewing Experts

Edited by

Alexander Bogner
Austrian Academy of Sciences, Austria

Beate Littig
Institute of Advanced Studies, Austria

and

Wolfgang Menz
Institute for Social Science Research, Germany

Research Methods Series

General Editors: Bernhard Kittel, Professor of Social Science Methodology, Department of Social Sciences, Carl von Ossietzky Universität Oldenburg, Germany and Benoit Rihoux, Professor of Political Science, Université catholique de Louvain (UCL), Belgium.

In association with the European Consortium for Political Research (ECPR), Palgrave Macmillan is delighted to announce the launch of a new book series dedicated to producing cutting-edge titles in Research Methods. While political science currently tends to import methods developed in neighbouring disciplines, the series contributes to developing a methodological apparatus focusing on those methods which are appropriate in dealing with the specific research problems of the discipline.

The series provides students and scholars with state-of-the-art scholarship on methodology, methods and techniques. It comprises innovative and intellectually rigorous monographs and edited collections which bridge schools of thought and cross the boundaries of conventional approaches. The series covers both empirical-analytical and interpretive approaches, micro and macro studies, and quantitative and qualitative methods.

Titles include:

Alexander Bogner, Beate Littig and Wolfgang Menz (editors)
INTERVIEWING EXPERTS

Audie Klutz and Deepa Prakash (editors)
QUALITATIVE METHODS IN INTERNATIONAL RELATIONS
A Pluralist Guide

Lane Kenworthy and Alexander Hicks (editors)
METHOD AND SUBSTANCE IN MACROCOMPARATIVE ANALYSIS

Research Methods Series

Series Standing Order ISBN 978-0230-20679-3 hardcover

(outside North America only)

You can receive future titles in this series as they are published by placing a standing order. Please contact your bookseller or, in case of difficulty, write to us at the address below with your name and address, the title of the series and one of the ISBNs quoted above.

Customer Services Department, Macmillan Distribution Ltd, Houndmills, Basingstoke, Hampshire RG21 6XS, England

palgrave
macmillan
Contents

List of Illustrations vii
Notes on Contributors viii

Introduction: Expert Interviews – An Introduction to a New Methodological Debate
Alexander Bogner, Beate Littig and Wolfgang Menz 1

Part I Theoretical Concepts:
Methodology of Expert Interviews

1 The Expert Interview and Changes in Knowledge Production
   Michael Meuser and Ulrike Nagel 17

2 The Theory-Generating Expert Interview: Epistemological Interest, Forms of Knowledge, Interaction
   Alexander Bogner and Wolfgang Menz 43

3 At Eye Level: The Expert Interview – a Talk between Expert and Quasi-expert
   Michaela Pfadenhauer 81

4 Interviewing the Elite – Interviewing Experts: Is There a Difference?
   Beate Littig 98

Part II Methodological Practice: Generating Data

5 On Interviewing "Good" and "Bad" Experts
   Jochen Gläser and Grit Laudel 117

6 Interviewing Experts in Political Science:
   A Reflection on Gender and Policy Effects Based on Secondary Analysis
   Gabriele Abels and Maria Behrens 138

7 Expert Interviews on the Telephone: A Difficult Undertaking
   Gabriela B. Christmann 157

8 Expert versus Researcher: Ethical Considerations in the Process of Bargaining a Study
   Vaida Obeléne 184
Part III - Fields of Application:
Applications of Expert Interviews in Different Fields of Research

9) How to Interview Managers? Methodical and Methodological Aspects of Expert Interviews as a Qualitative Method in Empirical Social Research
   Rainer Trinczek

10) Expert Interviews in Interpretive Organizational Research
    Ulrike Froschauer and Manfred Luenger

11) Between Scientific Standards and Claims to Efficiency: Expert Interviews in Programme Evaluation
    Angela Wroblewski and Andrea Leitner

12) The Delphi Method: Eliciting Experts' Knowledge in Technology Foresight
    Georg Aichholzer

Index

275

Table

2.1
5.1
5.2
6.1
12.1
12.2

Figure

4.1
5.1
5.2
7.1