1 Management consultancy viewed from economic and sociological perspectives

The literature on management consulting

Only since the 1990s has management consultancy prompted a great deal of attention in management research. Until then little had been written on this service sector, probably because it was not yet recognized as a mainstay in the economy. Management research, organization studies, and industrial sociology had primarily concerned themselves with larger industries and corporations, and the management consulting business was still too small to be recognized as an industry with considerable influence. Only a few authors, for example Hagedorn (1955), Higdon (1969), and Havelock and Guskin (1971), had begun to recognize the role of consultants in the transmission of business techniques. Other early publications on management consulting were concerned with organizational development, a consulting approach to help clients help themselves (Schein 1969; Argyris 1970).

Throughout the 1980s publications in the sociology of professions (Stanback 1979; Stanback *et al.* 1981; Noyelle and Dutka 1988; and later Tordoir 1995) referred to management consulting as one of the service sectors toward which industrialized economies shift. It became recognized as an emerging profession in which formal professional qualification has given way to professional work independent of a formal professional background (Abbott 1988; Brint 1994). At about the same time, Greiner and Metzger (1983) wrote a first advisory book for consultants, and the International Labour Organization (Kubr 1986) issued the second edition of a landmark book on best practices in management consulting, to which prominent management scholars and practitioners contributed and which aimed to cover a broad range of aspects from both consulting and client perspectives.

Despite these advances in the 1980s, the number of studies on management consulting remained low in comparison to the growth of the literature in the subsequent decade. Presumably it was assumed that not much could be added to the established view of consultants as transmitters of business techniques and carriers of organizational change methods. Not even the history of management consulting as a service sector and profession (McKenna 1995, 2001, 2006; Kipping 1996, 1997, 1999) was available to the scientific community before the 1990s. Only in the first half of the 1990s, following the rapid growth in the industry, did the significance and influence of management consulting become more recognized in the academic literature. Globally active consulting firms had achieved a high level of visibility, and management scholars could no longer ignore the influence of these firms on management knowledge, decisions, and practices. In the 1990s a large number of books appeared on the subject, oriented toward the markets for practitioners (e.g. Maister 1993; Kubr 1996), for MBA graduates applying to major consulting firms (e.g. Wet Feet Press 1996; Wickham 1999), or for those interested in starting their own consulting business (e.g. Kishel and Kishel 1996; Biech 1999).

At about the same time, a growing number of popular books on the potential dangers of hiring consultants appeared on the book market. These were mainly written by journalists or former consultants and had suggestive titles such as *The Inside Story* (Rassam and Oates 1992), *Dangerous Company* (O'Shea and Madigan 1997), or *Consulting Demons* (Pinault 2000). Even the Dilbert comics ridiculed consultants as shallow advisors. In this high tide of consulting bashing, well-known management scholars joined the ranks of those warning of *Flawed Advice* (Argyris 2000). Indeed, one of the salient characteristics of the consulting literature has been, and continues to be, that both journalists and academic commentators tend to have strong feelings about the business, considering consultants to be anywhere in a broad spectrum from shallow charlatans to modern carriers of economic growth.

Based on these images of the business, one can broadly distinguish between a functionalist and a critical view on consulting. The functionalist view sees consulting firms as carriers and transmitters of management knowledge. For example, Bessant and Rush (1995) distinguish between two knowledge-based roles for consultants: an intermediary one that supports clients' acquisition of knowledge and technological developments; and a capability-building one that supports clients' adoption and implementation of changes. Along this line, many authors have pointed out that consulting firms possess knowledge about analytical procedures which enables them to provide a variety of services and tasks that clients cannot perform on their own (Starbuck 1992; Moore and Birkinshaw 1998; Morris and Empson 1998; Sarvary 1999; Werr *et al.* 1997; Werr 1999, 2002; Armbrüster and Kipping 2002). Traditional organizations are assumed not to have the human resources, analytical skills, and procedural potential, with the result that taking management consultants into service has become a matter of course rather than an exceptional case, as it was some decades ago (Alvesson 1995; Faust 2002; Suddaby and Greenwood 2001). This perspective will be taken up in chapter 2 and integrated into a transaction cost perspective.

The functionalist view also points out other features of large consulting firms: the worldwide representation, the familiarity with a wide variety of industrial sectors, and the "one-firm" governance concept (for details, see chapter 8). These features ensure that consulting firms can obtain knowledge from a large variety of sources and, potentially, apply experiences gained in other industrial sectors or parts of the world. From this perspective, the methods to generate data and information outside and within the client organization constitute the primary driver of the consulting business and its growth. The recruitment of talented personnel, an extraordinary work ethic, and the strong commitment to an achievement culture represent a fundamental aspect of their performance and of the demand for their services. From the functionalist perspective, systematic knowledge management allows consulting firms to stay up to date with industry practices and market information, and it also enables them to distribute knowledge resources in a manner unequaled by conventional organizations (Larsen 2001; Hansen 1999, 2002; Hansen *et al.* 1999; Hansen and Haas 2001). I shall come back to these arguments in the transaction cost approach to consulting in chapters 2 and 8.

The critical literature on consulting does not necessarily doubt the usefulness of consulting for clients, but argues that the view that "consultants are experts and provide knowledge and analyses to clients for a fee" is too narrow to grasp what is going on in consulting projects (Clark and Fincham 2002). For example, Abrahamson (1996), Kieser (2002) and Ernst and Kieser (2002) refer to the faddish character of many management activities and argue that, among others, consulting firms have an economic interest in the up- and downswings of management concepts and substantially contribute to fashion setting. Berglund and Werr (2000) point to consultants' communicative flexibility, for example in their use of rationality and pragmatism myths to legitimate their approaches. Benders *et al*. (1998) have done empirical work in this context, finding that consultants use the term "business process reengineering" for a large variety of services that have often little to do with Hammer and Champy's (1993) original call for radical changes. Benders *et al*. (1998) argue that consultants separate the label from the contents of this management concept and create a sense of urgency by using a particular term without relating project contents to it. Similarly, Fincham (1995) argues that, in particular, business reengineering is constructed and marketed as a saleable commodity in order to meet the needs of the "managerial consumer." Ernst and Kieser (2002) and Kieser (2002) draw on these ideas to suggest that the circulation of management concepts and fashions contributes to managerial insecurity and fuels the demand for consulting services.

In a micropolitical view of consulting, Jackall (1988: 140–4) argues that consultants often trade in the troubles between the internal factions of a client organization, and that consultants often have to work on the problem as defined rather than develop a solution autonomously. As in an earlier approach by Moore (1984), client firms are not conceptualized as organizations as a whole, but as consisting of competing actors and groups. Using IT consulting as an example, Bloomfield and Danieli (1995) argue that the socio-political skills of consultants are indissoluble from their technical expertise, because technology cannot be separated from its communicative representation and thus from vested interests within a client firm. During the elaboration and implementation of advice, consultants and clients mobilize discursive and symbolic resources, which render it impossible to conduct consulting without any micropolitical involvement (see also Bloomfield and Best 1992). As with the other approaches, the micropolitical view draws on the insight that consultancy services are intangible and that their commercial impact is difficult to evaluate. But, rather than focusing on the consequential market mechanisms, the critical perspective on consultancy looks at the ways in which consulting assignments and client–consultant interactions are open to distortion.

In this context, Czarniawska-Joerges (1990) holds that the use of metaphors and labels that are new to the client organization can give meaning to situations and engender action through sense making. Seen from this perspective, the communicative resources of consultants provide some potential to obfuscate issues, to interpret situations for vested interests, or to manipulate definitions of success and failure. For Alvesson (1993), the point of departure is the uncertain character of all types of knowledge, even scientific knowledge. He argues that knowledge work needs to be viewed in the context of institutionalized myths of rationality, since there is no objectively determinable knowledge. Claims of knowledge, and therefore of communicative performance, may move into the foreground of this business, as credible stories about the world need to be delivered. The work of Clark (1995) has been influential in this respect. Given the lack of objective criteria for quality assessment, he argues, convincing clients of consulting quality requires considerable communicative skills and thus promotes consultants' impression management and rhetorical abilities. Along these lines, Clark and Salaman suggest viewing management consultants as "systems of persuasion creating compelling images which persuade clients of their quality and work" (Clark and Salaman 1998: 18).

In summary, the critical view argues that consulting results and project achievements are too problematic to be sufficiently theorized in terms of knowledge transfer. Authors in this paradigm point to the contestable nature of consulting knowledge, to the involvement of consultants in vested interests in client organizations, and to the potentially flexible mode of "consultancy speak." In so doing, they are expressing much of the concern, or even distaste, of an academic research community regarding consultants (March 1991), contributing to a more emancipated comprehension of the business. This critical take on consultancy will be taken up in chapter 4.

Theories used in this book

Publications of the above two types, the functionalist and the critical views, today characterize the literature on consultancy and have considerably advanced our knowledge of the industry and its mechanisms. Nevertheless, to date both are beset with limitations. The functionalist view lacks a systematic outline of why clients have increasingly externalized management services and continue to do so, and the critical view lacks an acknowledgment of economic processes and clients' rational deliberations. More precisely, the functionalist view presents useful lists or outlines of the economic role of consulting firms, but it lacks an analytical grounding. Neither theoretically nor empirically does it engage with the question of why client firms do not perform the services themselves or hire experts as employees rather than making use of external consultancies. It has not delved into the question of how clients gain quality certainty or why they hire a particular consultancy in preference to another, and a more theoretical analysis and elaboration suggests itself.

For its part, the critical view exhibits a limitation that is at least equally serious. As Salaman (2002) points out, it is preoccupied with consultants' truth claims, with consultants' supposedly unscientific approaches, and with an ostensibly dark side to consultancy. It either focuses on management fashions that clients supposedly fall for – which represents an oversocialized conception of the consulting market, to use Granovetter's (1985) term – or it portrays consultants as opportunistic agents who exploit clients' lack of quality certainty – which represents an undersocialized conception of management consulting. In some cases, the critical approach mixes over- and undersocialized views by portraying clients as somewhat retarded victims of both opportunistic consultants and mesmerizing management fads. This way, it has no concept of situations in which clients know exactly what they are doing when they hire consultants, and of conditions in which social ties and reputation effects preclude opportunistic action by consultants. Much of the literature from the critical camp seems based on an anti-consulting attitude, and scholars reproduce and reinforce their attitude in their research. The neglect, or even denial, of client prudence and economic deliberations is reminiscent of what W. O. Coleman (2002) has recently pointed out as anti-economics. I shall take up this discussion in chapter 4 and in the conclusion.

Sociological neoinstitutionalism

The only theory that the previous literature on consultancy has systematically drawn on is sociological neoinstitutionalism. For example, many articles in the volumes edited by Sahlin-Andersson and

Engwall (2002) and Kipping and Engwall (2002) draw on Meyer and Rowan (1977), DiMaggio and Powell (1983), Powell and DiMaggio (1991), or Tolbert and Zucker (1996). Sociological neoinstitutionalism is based on the argument that it is belief in the efficiency of particular practices or solutions, rather than any proven efficiency, that determines or influences economic action. According to this view, legitimacy toward the organizational environment rather than technical efficiency represents the core of organizing. If the efficiency or efficacy of organizational innovations or management ideas cannot be objectively evaluated, then they are oriented toward what the environment or decision-makers themselves believe to be efficient or effective. This leads to a number of effects – such as the institutionalization of management ideas – that are deemed efficient but are not necessarily so, or to pressure on organizations to adopt the same practices or structures as other firms (isomorphism) in order to gain legitimacy. Issues such as the legitimacy of organizational structures, the enforceability of change processes, and the validation of management decisions have taken center stage in the literature on consultancy (Sahlin-Andersson and Engwall 2002; Kipping and Engwall 2002; Alvesson 1993, 2004). The large and renowned consultancies in particular have duly been described as carriers not only of knowledge but also of legitimacy, as their analyses and reputation validate management decisions.

The diffusion of management concepts and innovations also touches upon elements of isomorphism in the neoinstitutional sense. If the efficiency and effectiveness of change initiatives or innovations often remain uncertain, then organizational decisions are frequently – on a normative or mimetic basis – oriented toward the behavior of other organizations. If a number of firms adopt a particular practice or innovation, then this is taken as signifying that these practices or innovations generate improvements. Even if it remains impossible to determine with certainty whether an innovation triggers progress or more efficient operations, a firm at least puts itself on equal footing with other firms if it adopts the same practices, and for this it often needs agents of change (such as consultants) as transmitters. Observations of McKinsey interventions, for example, have given rise to one of the founding publications of neoinstitutional theory, the article by DiMaggio and Powell (1983), which was based on the two authors' observation that McKinsey advice led to a number of isomorphic changes in public- and private-sector organizations.

Sociological neoinstitutionalism has been somewhat appropriated by the critical view on consultancy, as the theory seems to fit nicely into the critical camp's doubts about efficient outcomes from consulting assignments. However, the theory does not lend itself fully to the critical view. In fact, it has some elements of functionalism. For example, consultants as traders of legitimacy provide a service to a client *even if* their solution is similar to others, because it puts the consulted firm on a par with the others. Moreover, the sheer otherness of consultants in relation to client firms plays a central role in their ability to provide advice and gain legitimacy for it (Meyer 1996). And, as a central point, in their article on the institutional conditions for diffusion (of innovations, management practices, etc.), Strang and Meyer (1993) argue that any process of diffusion is accompanied or even preceded by a process of institutionalization. That is, before anything can disseminate as an idea or practice, it must be conceptualized and commodified as a term and concept, for only a communicatively transferable concept or explicit theory stands a chance of diffusing within or between professional groups. Consultants represent interpreters and theorists of individual cases and events. They often frame ambiguous information in new terms and theories, and thus develop and sharpen an interpretive consciousness within the client firm. Only this preceding theorization and term-building process enables an idea to diffuse. And, again, it is especially those consulting firms with a high public reputation that play a part in this process.

Signaling theory

The application of sociological neoinstitutionalism to management consultancy has been a useful and important advance, as it has highlighted the role of consultants in legitimation processes and in the communicative framing that precedes the diffusion of management concepts. Nevertheless, relying solely on sociological neoinstitutionalism may narrow the focus on societal norms and divert researchers from looking at the deliberative processes of individuals. Although sociological neoinstitutionalism acknowledges the possibility of different degrees of deliberation in economic action (Meyer and Rowan 1977; DiMaggio and Powell 1983), the question of the conscious behavior of economic actors represents the Achilles heel of this theory. As DiMaggio (1988: 9) observes, "[s]elf interested behavior tend[s] to be smuggled into institutional arguments rather than theorized explicitly." Sociological neoinstitutionalism has been developed to model the influence of norms on economic action, but it has difficulties with modeling autonomous action in the context of norms that economic actors are aware of. DiMaggio's (1988) distinction between institutionalization as a process and as an actual state then allows us to conceive of individual action at least in processes of institutionalization (see, in this context, Tolbert and Zucker 1996 and Barley and Tolbert 1997). If we take into account the possibility that clients are experienced and knowledgeable executives who can reflect on norms and act deliberately, then sociological neoinstitutionalism meets its limits and other theories suggest themselves.

In particular, economic signaling theory (Spence 1973, 1974, 1976) models deliberate signaling processes in the context of known norms. Signaling theory argues that, in markets of credence goods and quality uncertainty, providers invest in product or service features that signal status, quality, and reliability. Spence models graduate education (essentially, the reputation that different kinds of education involve) as a signal for graduates' future productivities. At the center of attention are the costs of signaling (e.g. for graduates on the job market the costs of education such as loans and household credit, and the effort put into attaining the degree), the effects of signaling (type of job, salary, promotions of the hired employee), and the incentive structures to invest in signals. If a provider cannot prove the quality of the outcome prior to purchase, and not even for some period after purchase, then he resorts to proving input factors as an indicator for the quality of the outcome. Signals such as certificates concerning educational background reduce the information asymmetry between supply (graduate) and demand (employers) of labor. Spence's central point is that a good education works as an efficient mechanism to signal a graduate's future productivity because, for someone with lower future productivity, it would be much more costly (investments, efforts) to attain a renowned degree. A conceptually simple but methodologically unfeasible test of signaling theory would be, for example, to gather people of identical ability and randomly assign some of them a degree certificate. If those with the certificate later earn more, then signaling theory would be supported.

There is one fundamental difference between economic signaling theory and sociological neoinstitutionalism. The former assumes that the signaling mechanism works as an efficient device to connect supply and demand, the latter looks at deviations from economic efficiency that legitimacy-seeking behavior brings about. In other words, signaling theory assumes that the market clears efficiently and conceptualizes how this comes about by signaling mechanisms. This assumption of efficiency may appear absurd to sociological neoinstitutionalists, because they observe economic action oriented toward norms and anticipated expectations independent of or detrimental to economic efficiency. Indeed, the explanation of economic actions in cases where efficiency remains unclear is the main purpose of the theory.

Nevertheless, the two theories have two important aspects in common. First, both view the essence of economic behavior in aspects external to the immediate exchange relationship, such as the status of education at prestigious colleges/universities or the status of particular concepts of organizational structure. In other words, both focus on the orientation of economic behavior toward the norms within which exchange partners act, rather than toward the immediate features of the exchange partners. The second commonality is that both theories imply a decoupling of reputation from the actual quality of a service. For sociological neoinstitutionalists, the legitimacy effect is decoupled from the economic quality of a decision (e.g. regarding organizational structure). Alternatively, an economically positive effect arises as a result of the gained legitimacy rather than from any intrinsic economic quality of the decision. Firms make particular decisions not because they have proven economic effects but because the environment considers them useful. Signaling theory, too, relies on the assumed rather than the actual quality of education. That is, a graduate from a college of high reputation may have undergone a worse preparation for a job than someone from an unknown college. Nevertheless, the graduate from the high-reputation college is rightly assumed to have a higher future productivity. This is because those individuals with a high future productivity independent of the education have less costly access to colleges of high reputation. Thus the signaling mechanism works irrespective of the actual quality of the education.

Important for our purposes is the notion that whether the behavior of market participants leads to efficient or inefficient outcomes cannot be assumed a priori, hence there can be no prior nonnormative preference for either of the two theories. Rather, the essence is to compare the theories with regard to individual phenomena.