

Epistemics at Work

The Theory of Mind in Principal-Agent Relations

Linder, Stefan; Foss, Nicolai J.; Stea, Diego

Document Version

Final published version

Publication date:

2014

License

CC BY-NC-ND

Citation for published version (APA):

Linder, S., Foss, N. J., & Stea, D. (2014). *Epistemics at Work: The Theory of Mind in Principal-Agent Relations*. Institut for Strategic Management and Globalization. SMG Working Paper No. 3/2014

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025



**EPISTEMICS AT WORK:
THE THEORY OF MIND IN PRINCIPAL-AGENT RELATIONS**

Stefan Linder
Nicolai J. Foss
Diego Stea

SMG WP 3/2014

February, 2014

SMG Working Paper No. 3/2014
February, 2014
ISBN: 978-87-91815-95-9

Department of Strategic Management and Globalization
Copenhagen Business School
Kilen, Kilevej 14A
2000 Frederiksberg
Denmark
www.cbs.dk/smg

EPISTEMICS AT WORK:
THE THEORY OF MIND IN PRINCIPAL-AGENT RELATIONS

Stefan Linder
ESSEC Business School
Avenue Bernard Hirsch, BP 50105
95021 Cergy-Pontoise (Paris) / France
linder@essec.fr

Nicolai J. Foss
Department of Strategic Management and Globalization
Copenhagen Business School
Kilevej 14A, 2nd fl.
2000 Frederiksberg / Denmark
njf.smg@cbs.dk

Diego Stea
Department of Strategic Management and Globalization
Copenhagen Business School
Kilevej 14A, 2nd fl.
2000 Frederiksberg / Denmark
njf.smg@cbs.dk

February 24, 2014

Keywords: Agency theory, principal-agent relations, epistemic assumptions,
theory of mind.

JEL codes: D23, D82, D86, M5

EPISTEMICS AT WORK:
THE THEORY OF MIND IN PRINCIPAL-AGENT RELATIONS

Abstract

Agency theory studies the impact of and remedies to asymmetrically distributed information in principal-agent relations. Yet, it does so in a surprisingly binary manner: It assumes the principal to be perfectly knowledgeable of some pieces of information (such as the agent's risk aversion), while others (such as the agent's true effort exerted) are considered to be perfectly private information of the agent. Agency theory thus makes very asymmetrical assumptions about the knowledge of principals and agents, largely neglecting the human capacity for interpersonal sense-making. This chapter explores the implications of instilling agency theory with a more realistic account of the human capacity to read other people's desires, intentions, knowledge, and beliefs—that is, to have a theory of someone else's mind.

I. Introduction

Agency theory has profoundly influenced our understanding of a wide range of principal-agent relationships within labor market and business contexts, as well as in sociology and political science research (e.g., Adams, 1996; Holmström, 1989; Lazear, 2000; Rose-Ackerman, 1975). The seemingly simple idea that agency problems can largely be remedied by relying more on performance-contingent incentives instead of fixed salaries, turned pay-for-performance into a sort of management mantra among practitioners in the late 1990s. Increasingly, however, the theory attracts criticism (e.g., Ferraro, Pfeffer, and Sutton, 2005; Heath, 1999; Perrow, 1990). In particular, many authors have criticized classic agency theory as painting a too dark picture of human nature by assuming agents to be self-interested utility maximizers, to always experience “disutility of labor,” and thus to engage in the behaviors desired by the principal only when motivated by extrinsic incentives.

Yet, as the basic agency model shows, conflicting interests among a principal and an agent only become problematic if the two parties differ in the information available to them (e.g., Laffont and Martimort, 2001; Milgrom and Roberts, 1992): Moral hazard is a concern only if the principal cannot observe the agent’s actions. Similarly, the hidden characteristics problem arises from the principal’s inability to identify (*ex ante*) the true characteristics of the agent. Hence, the implicit or explicit assumptions about what the two parties know, how they process information, and how they got to know what they know in the first place impact the insights that one gains from agency theory as much as the assumptions about self-interest and (extrinsic) motivation. However, whereas the motivational foundations of agency theory have sparked significant debate, the cognitive and epistemic assumptions of the theory have not attracted much attention yet (Foss and Stea, 2014; notable exceptions are Hendry, 2002, and Pepper and Gore, 2013).

The present chapter sheds new light on the epistemic assumptions of agency theory and explores the implications and potential value of instilling agency theory with a realistic

account of the human ability to infer another person's desires, intentions, knowledge, and beliefs—that is, to develop a “theory of mind” (Frith and Frith, 2003; Singer and Fehr, 2005).

We begin with a sketch of agency theory with a particular reference to the hidden action problem, which allows us to highlight the theory's epistemic assumptions. We then discuss the human capacity to develop a theory of mind, and the ways in which such a concept can further our understanding of principal-agent relations. A concluding discussion closes the chapter.

II. Agency Theory and its Epistemic and Cognitive Assumptions

a. The Starting Point of Agency Theory: Delegation

Starting point of standard agency theory is the common situation in which one person (the principal) would like to delegate some task to another person (the agent) – either because the principal has too high opportunity costs to carry out the task herself or because she lacks the knowledge of how to best perform that task (Jensen and Meckling, 1976; Laffont and Martimort, 2001). The principal thus first needs to recruit an agent (phase 1, or contracting phase). Subsequently, she can then have that agent carry out the task for her (phase 2, or post-contractual phase). Yet, the agent's interests are assumed to differ from those of the principal. Such differences in, for example, their outcome-type preferences (Holmström, 1979; Jensen and Smith, 1985), risk preferences (e.g., Grossman and Hart, 1983), or time preferences (e.g., Baker, Gibbons and Murphy, 1994), may imply that a self-interested, rational and utility maximizing agent “will not always act in the best interests of the principal” (Jensen and Meckling 1976: 5). As a consequence, the principal should observe the agent's acts and pay him based on these acts thereby providing the proper incentives for the agent to act in the manner desired by the principal (e.g., Jacobides and Croson, 2001).

However, a second, key aspect of agency theory is information asymmetry. Specifically, the agent is assumed to be better informed about her own characteristics, actions, and the state of nature than the principal. Agents may exploit this information asymmetry by

either trying to get recruited by the principal and to earn an income even though they lack the skills or abilities to properly perform the task (adverse selection, or hidden characteristic problem), or once recruited by withholding effort, or engaging in actions not valued by the principal (moral hazard problem).

In sum, conflicting interests and a principal facing an informational disadvantage give rise to the two problems studied within agency theory – “adverse selection” or “hidden characteristics” in phase 1 and “moral hazard” or “hidden action” in phase 2 (Grossman and Hart, 1983; Holmström, 1979). Agency theory aims at providing recommendations about how principals can minimize these problems.

b. Standard Agency Theory: Second-best Solutions

The wish to reap the benefits of delegating a task to an agent and the assumed conflict of interest between the two parties thus set the stage for the principal contracting with the agent about a wage W . In the standard agency model (see e.g., Milgrom and Roberts, 1992, pp. 222n), this wage W consists of some fixed salary component $W_0 \geq 0$ and a proportion β of the output z generated by the agent (i.e. result attained). Agency literature typically refers to this proportion β (where $0 \leq \beta \leq 1$) as the “incentive intensity” of the contract.

The agent in turn decides on whether to accept the contract offered by the principal (in phase 1) and subsequently on the actions she takes (in phase 2). For convenience, most agency theorists think of the agent’s actions as his effort e , that is, alternative actions correspond to different effort levels (e.g., Holmström, 1979). Thus, the terms “effort” and “action” are often used interchangeably. The model further then assumes the agent’s choice of effort e to affect – together with some stochastic influence from the environment θ – the outcome z attained, and hence, the payoffs of both parties (Grossman and Hart, 1983; Holmström and Milgrom, 1991; Jensen and Meckling, 1976; Ross, 1973). For simplicity, the basic models assume this external influence to be standard-normally distributed with an expectation of zero. Moreover, agents are assumed to have exponentially growing disutility of

engaging in effort ($C(e)$) and to consider these “costs” when deciding on whether to accept a contract and on choosing their effort level e (Grossman and Hart, 1983). Principals therefore need to consider the agent’s costs of acting in a manner functional for maximizing the principal’s payoffs ($C''(e)$) when contracting with the agent over W , that is: when solving their own optimization problem concerning which β they should offer to the agent.

They also need to consider that contracting with the agent over a proportion $\beta > 0$ of the output z , means exposing the agent to income risk since z depends on both, the agent’s efforts e and the stochastic influence from the environment θ over which the agent (assumably) has no control. A risk-averse agent will therefore ask for a risk premium in order to accept a contract offered by the principal based on the risky result z . This risk premium does not give the agent extra utility and merely reduces the net payoffs the principal receives. Thus, the larger the risk premium, the smaller the total value created in the relation (i.e., the sum of the monetary equivalents of the utilities of the principal and the agent). The size of the risk-premium depends on both the degree of risk-aversion of the agent r_A and the riskiness of the wage (i.e., the variance in the wage due to the environment’s impact ($Var(W) = \beta^2 Var(\theta)$)). Thus, β has to solve two tasks: providing incentives to the agent to work hard and insurance against risk. Agency theory therefore calls for principals to consider their agent’s degree of risk aversion r_A in designing optimal contracts. Hence, standard agency theory assumes the principal and the agent to know not only the distribution and the expectation of the stochastic influence from the environment θ and the agent’s costs $C''(e)$ of acting in a manner functional for maximizing the principal’s payoffs, but also the agent’s risk aversion r_A .

In contrast to the principal’s strong knowledge about these factors, the true effort e exerted by the agent and the actual state of the environment θ are assumed to be only known to the agent; the principal is considered to be able to observe only the result z caused by both e and θ . The principal’s inability to disentangle e from θ – that is, the information asymmetry

between the principal and the agent for these two variables (and only these!) – then is seen as implying that any solution is inefficient (“second-best”) as compared to a hypothetical first-best solution in which the (assumably risk neutral) principal would bear all the risk and the (assumably risk averse) agent would be paid on his/her efforts e only.

Based on this assumed distribution of information among the two parties, a few mathematical optimization steps allow finding the optimal β that leads to such a second best solution to the moral hazard (i.e. hidden action) problem. The starting point is to maximize the total certainty equivalent for both parties to the optimal incentive intensity (denoted as β^*). A couple of mathematical steps yield the following expression $\beta^* = P'(e) / [1 + r_A \cdot \text{Var}(\theta) \cdot C''(e)]$ (for the details of the a mathematical derivation see e.g., Milgrom and Roberts, 1992: 222n; and for the original formulation, Holmström and Milgrom, 1991).

Standard agency theory thus suggests that the principal can address the hidden action problem by contracting with the agent over the optimal β^* , which depends on four factors: (1) The more sensitive the payoff to the principal on changes in the agent’s efforts or actions ($P'(e)$), the larger should be β^* ; (2) the stronger the agent’s risk aversion r_A , the lower the β^* ; (3) the larger the riskiness of the results obtained by the agent due to environmental influences $\text{Var}(\theta)$, the lower the incentive intensity; and (4) the stronger the agent’s discretion regarding the choice of activities ($\partial e / \partial \beta = 1 / C''(e)$), that is, the lower his/her costs of acting in a manner functional for maximizing the principal’s payoffs $C''(e)$, the larger the β^* .

Either independently or jointly with “moral hazard” (hidden action), the principal may also face “adverse selection” (hidden characteristics). That is, the principal may face an informational disadvantage with respect to the agent’s characteristics (beyond r_A and $C''(e)$), such as his knowledge, skills, and abilities. One cause for such a disadvantage is that the very reason for delegating a task is often that the principal lacks the knowledge, abilities or skill to perform the task herself (see e.g., Ross, 1973). Hence, the knowledge differences that explain

why gains from trade exist in the first place give rise to problems for the principal in judging the agent's true qualities. The situation is complicated by the fact that self-interested agents may deliberately hide their true qualities (e.g., a lack of appropriate knowledge or skills for conducting the task) or even falsify signals (e.g., faking the degrees they hold) in order to get a job and earn the related payoffs.

Choosing the right β^* may also help the principal with respect to reducing this adverse selection problem by triggering self-selection of skilled agents into the contract. Alternatively, the principal may engage in screening actions, or the agent may engage in signaling. Whereas the first out of these three remedies to the hidden characteristics problem aims at motivating agents to indirectly reveal their characteristics, the latter two are direct, explicit strategies aimed at reducing the informational disadvantage that the principal faces about the agent's characteristics (beyond r_A and $C''(e)$, obviously). All three strategies to tackle the adverse selection/hidden characteristics problem, however, are costly: either they imply paying the agent a risk-premium which is taken from the surplus or they cause costs for screening or signaling. Given these agency costs, all of these three routes lead only to a second-best solution.

c. The Silent Power: Agency Theory's Cognitive and Epistemic Assumptions

As the preceding outline of agency theory shows, the theory studies principal-agent relations in a setting of clearly specified bits of information that are available to either both the principal and the agent (such as r_A and $C''(e)$), or to the agent alone (like e and θ). In fact, it is the assumptions about what the parties know that are essential for whether an agency problem emerges. Only under the assumption of information asymmetry the conflicting interests become problematic. For this reason, the assumptions about what the two parties know, how they process information, and how they got to know what they know have just as much impact on the insights that can be gained from agency theory as do the assumptions about self-interest and (extrinsic) motivation underlying the theory.

Table 1 summarizes the information that standard agency theory assumes principals and agents to hold and to use in maximizing their utility. Many of these assumptions we already touched upon in the preceding section.

Table 1 about here

Already a brief glance at Table 1 suggests that the notion of agency theory building on asymmetrically distributed information would benefit from a qualification. In fact, agency theory highlights asymmetrically distributed information of few of the several variables that are considered while at the same time principal and agent are assumed to have perfect common knowledge of many other variables. For example, both the principal and the agent are assumed to know the set of possible actions that the agent can take (or should take from the principal's perspective)—that is, to know the full space of e . This, however, presupposes that the principal is more or less as knowledgeable about the task at hand as the agent – otherwise he would have troubles knowing all the possible actions related to, or necessary for the task. Such a setup, however, precludes studying principal-agent relations where the principal hires an agent because he exactly lacks the necessary knowledge to carry out a task himself. Hence, the fundamental question of what the principal should want the agent to do is (implicitly) assumed away in standard agency theory (Hendry, 2002). It thus starts – as opposed to transaction cost economics – from so called “complete contracts”. Whereas Ross (1973: 135) submits that under this assumption “the problem is considerably simplified but much of the interest does remain,” the assumption in standard agency theory that the principal knows the space of e hampers its usefulness for studying principal-agent relations that imply tasks where the principal faces a knowledge disadvantage – not only an information asymmetry – vis-à-vis the agent.

Unfortunately this type of principal-agent relations are increasingly important in business practice. Knowledge workers have become a key, if not dominant group of agents in most firms. Along the same line, studying entrepreneurship by definition becomes pointless in settings where the full space of options is already known and the future implies only probabilistic issues, but not uncertainty (Knight, 1921; see Foss and Klein, 2012, for a recent discussion on entrepreneurship and organizational economics). Thus, making many practical phenomena amenable to an analysis of principal-agent relations requires choosing a larger (or different) set of variables for which information is distributed asymmetrically than what is assumed in standard agency theory.

Given the epistemic assumptions of standard agency theory, it would be more accurate to speak of the theory as studying problems of asymmetric information on effort and agent type, rather than of problems of asymmetric information in general. Moreover, the theory is largely silent on why (or why not) the two parties have particular pieces of information—that is, on how the parties got to know what they know in the first place. For example, one may wonder why the principal is assumed to have perfect knowledge (i.e., why he has the same beliefs as the agent) about the expected value ($Exp(\theta)$) and distribution ($Var(\theta)$) of the states of nature (Ross, 1973), even though he is – unlike the agent – assumed to be unable to ascertain the true state of nature θ and thus unable to step-by-step update his potentially false prior beliefs to eventually arrive at shared beliefs with the agent about $Exp(\theta)$ and $Var(\theta)$. In sum, agency theory views most variables as common knowledge without further explaining how the individuals involved in the relation actually got that knowledge.

Further, agency theory assumes that the principal knows the agent to be fully rational. The agent, in turn, knows that the principal knows (that the agent knows) and so on (Foss and Stea, 2014; Robalino and Robson, 2012). Thus, agency theory in its ascription of rationality follows the “common knowledge” assumption underlying most modern game theory (Holler, 2002). Hence, whereas agency theory is sometimes portrayed as resting on foundations of

bounded rationality (e.g., Eisenhardt, 1989), it in fact assumes both parties to make full and utility-maximizing use of the information available (Hendry, 2002; Holmström, 1979; Holmström and Milgrom, 1991; Laffont and Martimort, 2001; Ross, 1973).

For some particular variables pertaining to a principal-agent relation – for example, the timing of the actions in the relations – assuming common knowledge is unproblematic. However, just like work in social-psychology and in experimental economics demonstrates that humans typically do not possess the cognitive apparatus needed to maximize expected utility in a Bayesian sense (except if decision situations are trivial) (Camerer, 1998; Kahneman and Tversky, 2000; Robalino and Robson, 2012), insights from evolutionary anthropology, neuroscience, (neuro-)economics, and social psychology on the “theory of mind” suggest that the strong assumptions about what the principal and the agent know, and what they know about each other are overly simplifying real-world principal agent relations (cf. Robalino and Robson, 2012).

In fact, the epistemic assumptions of agency theory on the one hand largely overestimate what principals and agents can know in reality. For example, in many real-life cases, principals when recruiting a new agent will suffer from much broader informational disadvantages than assumed in standard agency theory. Very often they will not know the agent's r_A nor $C''(e)$ or at least their beliefs about these variables will differ from the actual r_A and $C''(e)$. In other words, most real-life principals are unlikely to see so clearly into their agent's brains to ascertain the exact risk aversion of each of them. While they may try to form beliefs about these variables from observing their behavior, emotional reactions, stereotyping agents, or by projecting their own characteristics onto agents, chances are good that these beliefs differ from those held by the agent about himself.

On the other hand, however, the epistemic assumptions also underestimate the principal's ability to gain insights on some variables that he is explicitly assumed not to be able to know at all, such as the agent's type or actions. Yet, many of us (and certainly those

who have children) can probably recall situations in which we were unable to observe another person's actions or efforts, but where we were able to infer them – not from the result z – but from, for example, the person's non-verbal body language. Decoding those cues draws on the ability to develop a theory of another person's mind. This ability is disregarded in standard agency theory, and yet highly important for social interaction and interdependent decision making (De Melo, Carnevale, Read, and Gratch, 2014).

III. Theory of Mind

a. Defining 'Theory of Mind'

Scholars in various traditions have suggested that the capacity to infer others' desires, intentions, knowledge, and beliefs is crucial for interdependent decision making because the beliefs about others' states of mind bear an impact on one owns choice of behavior (e.g., Ames, Weber, and Zou, 2012; Aumann and Brandenburger, 1995; Galinsky et al., 2008; Gallagher and Frith, 2003; Singer and Fehr, 2005). Not surprisingly thus, research in areas as different as evolutionary anthropology, social psychology, neuroscience and neuro-economics has been investigating this capacity, its antecedents, and consequences under a plethora of different labels.

Premack and Woodruff (1978) coined the term “theory of mind” (ToM) to denote the ability to read the desires, intentions, knowledge and beliefs of other people. Others refer to this capacity as “mentalizing”, “mind reading”, or “cognitive perspective taking” (e.g., Bagozzi et al. 2013; Baron-Cohen, Leslie, and Frith, 1985; Frith and Frith, 2003; Singer and Fehr, 2005). Whereas subtle nuances in the conceptualization of this capacity and its related processes exist among the various researchers and scholarly traditions, there is a general agreement that the capacity to explain and predict other people's behavior by attributing to them independent mental states, such as beliefs and desires, allows to “allows an individual to anticipate the behavior and reactions of others” (Galinsky et al., 2008: 378) and underpins their ability to deceive, cooperate and empathize (Gallagher and Frith, 2003).

ToM is similar, yet distinct from concepts such as transactive memory, empathy, general sensemaking, and thinking. The concept of transactive memory shares with ToM that it involves an understanding of what others know. However, ToM not only relates to knowledge, but comprises an understanding of others' desires, intentions, and beliefs (Foss and Stea, 2014). Likewise, empathy differs from holding a ToM. It involves an emotional stance, whereas the notion of ToM exclusively refers to the cognitive theorizing about another individual's desires, intentions, knowledge, and beliefs (Bagozzi et al., 2003; Hein and Singer, 2008). A ToM also differs from the more general concept of sensemaking (Weick, 1995) in that it refers only to an understanding of another individual's mental states. Finally, forming and holding a ToM while being a form of thinking, clearly differs from the general term. A ToM implies thinking about particular elements – namely, desires, intentions, knowledge, and beliefs – in another person's mind (Singer and Fehr, 2005).

Being able to put oneself into the shoes of some other person has been associated with such individual characteristics as, for example, high self-esteem and low neuroticism (Davis, 1980; 1983). However, there is not yet a full account of the mechanisms that underlie the development of a theory of mind. What is clear, however, is that an individual who forms or holds a ToM needs the ability to “attribute mental states – desire, knowledge, belief, intent, etc. – and to interpret observed behaviour in terms of such states” (Robalino and Robson, 2012: 2224). Hence, it is essential to recognize that, for example, a person's belief might not correspond to reality, and that nevertheless “it is the belief, not the reality, that determines behaviour” (Gallagher and Frith, 2003: 77). Not all individuals are equally well equipped to recognize this difference. Thus, research suggests that little children who have not yet developed the ability to attribute mental states to others or individuals, have not yet a ToM. Likewise, research suggests that people high on autistic spectrum disorders seem to be characterized by a selective impairment in the theory of mind (Dietvorst et al., 2009; Gallagher and Frith, 2003).

b. How people mentalize

Individuals seem to rely on one or multiple routes for inferring others' desires, intentions, knowledge, and beliefs. Evidence by Ames et al. (2012) suggests that people rely systematically on projection and stereotyping to infer others' mental states. Besides using projection and stereotyping, humans seem adept at inferring others' intentions, desires, or beliefs based on such sources as simple body movements and facial expressions – in particular: emotion expressions, as well as elaborate sequences of actions that converge on desired outcomes (Ames, Weber, and Zou, 2012; De Melo, Carnevale, Read, and Gratch, 2014). For example, De Melo et al. (2014) recently experimentally demonstrated that people make inferences about others' mental states from emotion expressions in interdependent decision making. In other words, emotion expressions in fact transmit information that helps inferring the intentions, desires, and beliefs of one's counterpart (De Melo et al., 2014; Van Kleef et al. 2010). Agency theory's categorical assumption that the principal cannot ascertain whether a good result z is due to an agent's high efforts e or caused by a favorable state of nature, therefore, implicitly assumes a principal that is completely unable to read subtle signs of body language, emotion expressions or the like or, in other words, to put himself into the agent's shoes.

More specifically, when looking at the epistemic assumptions underlying standard agency theory, principals seem to be characterized by a selective and highly asymmetrical impairment in ToM: They can perfectly read agents' $r_A, C''(e)$, as well as the expected values of nature and their distribution from their agents' minds, but miserably fail in mentalizing about the agent's actions e . Given this selective impairment, one might be tempted to ascribe standard agency theory's principals a significant level of autistic disorder.

IV. The Role of Mentalizing in Principal-Agent Relations

As our discussion shows, standard agency theory in some important aspects over-estimates the epistemic powers of principals (and agents), whereas it under-estimates them in others.

Considering the impact of a ToM in principal-agent relations requires a refinement of the canonical model described in section II. Given space constraints, we cannot lay out a full and nicely honed revised theory here, but only illustrate some exemplary implications of such a theory and its recommendations on how to handle agency problems. We do so by discussing the implications of considering the (limited) human capacity to mentalize first in the contracting phase (phase 1, as denoted it in section II), and then in the post-contractual phase (phase 2).

a. The impact of a ToM in the contracting phase

Chng, Rodgers, Shih, and Song (2012) recently found that executive core self-evaluation influences the effects of compensation schemes. This finding lends support to Wowak and Hambrick (2010), who emphasized the role of executives' characteristics and the need to consider these aspects in designing incentive systems for them. Assessing such characteristics as core self-evaluations, however, requires a principal to make full use of the ability to read (potential) agents' emotion displays or sequences of actions. The principal's theory of the agent's mind is thus likely to have a significant impact on the design of the contracts that the principal offers to the agent.

Whereas standard agency theory assumes the principal to know both r_A and $C''(e)$, principals with imperfect mentalizing capacity may misestimate these values. This may lead, for example, to the principal believing the agent to incur lower than the true costs $C''(e)^p < C''(e)$. If that is the case, the principal will thus offer the agent too little compensation for exerting effort (i.e. a $\beta < \beta^*$). This would reduce the agent's willingness to exert effort (since $C(e(\beta))$ and, in turn, her own payoffs $P(e(\beta))$). Alternatively, it might be that a principal with an imperfect theory of the agent's mind overestimates r_A (that is: $(r_A)^p > r_A$) and, hence, shies away from offering the optimal β^* suited for r_A and goes for a $\beta < \beta^*$. This would entail forgoing some of the incentive effects of variable compensation and again lowering the

payoffs to the principal. Since the principal with an imperfect theory of mind might misestimate both r_A and $C''(e)$, the contract he offers to the agent might be quite different from the second-best solution suggested by standard agency theory. Obviously, the principal might equally well offer a $\beta > \beta^*$ if he misreads the agent's r_A and/or $C''(e)$. For example, if he assumes the agent to incur higher costs than is true, the principal may offer the agent a contract with “excess compensation” for exerting effort ($\beta > \beta^*$). This in turn, *ceteris paribus*, implies exposing the agent to higher risk than is optimal (that is, too much incentive but too little insurance). This may lead to a situation where the total value created in the principal-agent relationship diminishes as compared to the second-best solution. Some payoffs that would accrue to the principal are shifted to the agent who, however, does not gain in utility as this shift is eaten up by the increase in the risk premium he needs to get the same utility – that is, to accept the risky compensation scheme. Obviously, even with the principal starting with false beliefs about r_A and $C''(e)$, the principal and agent might nevertheless arrive at the second-best solution of standard agency theory if by accident the effects of the misestimates about r_A and $C(e)$ cancel each other out. Yet, this is likely to be rather an exception, and hence, instilling agency theory with an account of imperfect mentalizing capacities is likely to imply rather a third-best solution than a second-best one.

Likewise, the agent's capacity to read the principals mind might equally matter in the contracting phase. In reality, few labor markets work the way assumed in standard agency theory with the proposing contracts on a take-it-or-leave-it basis. Rather, some bargaining between the principal and the agent takes place. Thus, the principal needs to consider this bargaining process when proposing contracts. The agent's theory of the principal's mind is likely to play a key part in this process. For example, an agent with a highly imperfect ToM may tend to accept a contract that the principal deemed still preliminary (that is, to be further refined in the negotiation). On the other hand, an agent with an accurate theory of the principal's mind may be able to recognize that the principal wants him to act as an

entrepreneur, and thus suggest or accept more subjective performance appraisal practices given the difficulty of specifying objective criteria (cf. Baker, Gibbons and Murphy, 1994). An agent with low mentalizing capacity, in contrast, may misread the principal's intentions and maybe assume the principal to aim for a contract that allows the principal to opportunistically renegotiate the payment scheme ex post (Prendergast and Topel, 1996). Moreover, agents who have troubles reading potential principals' minds may end up signing a contract with a principal who shirks, cheats, opportunistically seizes perquisites for him, or exploits her agents (see Shapiro 2005).

In sum, considering the agent's ToM allows accounting for the real-world phenomenon of agents signing contracts with opportunistic principals. This may help in addressing a point of criticism voiced against agency theory by Perrow (1990) and others, namely that the theory does not fully capture both the principal's and the agent's side, and tends to overlook that there are not only different types of agents, but also different types of principals.

Returning to the case of the principal misreading the agent's $C''(e)$ and therefore offering the agent a too high compensation for exerting the task, an agent with a very imperfect theory of the principal's mind may in turn misread this signal. Instead of realizing that the principal simply misjudged the agent's costs, that agent may for example interpret the higher compensation offered (i.e. the $\beta > \beta^*$) as a signal that the task is particularly cumbersome, boring, or unpleasant in other forms. In other words, the agent might erroneously assume that the principal knows the effort necessary for carrying out the task better than he himself does and that the $\beta > \beta^*$ shows that he should increase his own estimate of $C''(e)$. This might imply that the actual effort exerted remains more or less the same – that is, that $\beta > \beta^*$ will not result in an increased output z . Thus, imperfect mentalizing on the principal's and the agent's sides at the same time might lead to a situation that mirrors

the “overjustification effect” or “crowding-out” social-psychologists ascribe to extrinsic incentives on task-intrinsic motivation (Deci, 1971; Lepper, Greene, and Nisbett, 1973). Considering the ToM of both principal and agent may thus provide a way to reconcile formal agency theory with extant experimental evidence accumulated in social-psychology (e.g., Deci, Ryan, and Koester, 1999), and increasingly in economic and management research, too (e.g., Fessler, 2003; Kunz and Linder, 2012).

b. The impact of a ToM in the post-contractual phase

We have seen that standard agency theory assumes the principal to be unable to ascertain whether a good result z is due to the agent’s high effort level e , or rather to “luck” (θ). For this reason, the principal has to expose the agent to income related risk, and thus pay him a risk premium. Absent the principal’s ToM, the standard agency model applies. However, if the principal is assumed to have a theory of the agent’s mind, then that theory will allow reducing the informational asymmetry she faces and thus the problem of hidden action decreases. This obviously has important implications for the optimal measurement and incentive intensities. Ceteris paribus the principal can offer the agent a higher β as the principal’s theory of mind allows reducing the $Var(W)$ that the agent faces since the principal then is better able to tease out the agent’s efforts e when observing z . Therefore, the agent will agree to receive a lower risk premium than in the setting of standard agency theory. Alternatively, the principal could keep the β as it is, but lower his monitoring efforts, thereby again lowering overall agency costs.

The role of the principal’s ToM for principal-agent relations can also be illustrated by looking at the moral hazard problem and incentive provision in a multi-tasking setting. Here, standard agency theory suggests that agents working on tasks with multiple dimensions, some of which easy to measure (and hence to reward) and others difficult or impossible to assess by the principal, should receive only weak incentives or a flat salary (Holmström and Milgrom, 1991). However, in practice it is fairly easy to find examples of firms violating this

recommendation. Obviously, one way of explaining this phenomenon is to rely on fairness arguments. In other words, firms go for a variable compensation scheme even though agency theory would suggest differently simply because rewarding one agent group (that does multi-tasking) differently than another group (that is not multi-tasking) may trigger perceptions of procedural unfairness, inequality, and foster envy (Bewley, 1999; Pfeffer and Langton, 1993). Considering ToM may provide an alternative explanation: Given a multi-tasking situation, choosing a stronger incentive intensity than recommended by standard agency theory (i.e. a $\beta > \beta^*$) may also suggest that, while being unable to objectively assess the agent's performance on all dimensions, the principal's theory of the agent's mind allows her to infer the agent's true performance on all dimensions by only measuring some. In other words, the principal's ToM allows her to judge whether the agent "took shortcuts" to maximize the performance achieved on the measured parts.

Clearly, over time the agent might learn that the principal's ToM is not as high as the $\beta > \beta^*$ would require. In other words, if the agent "got away" with maximizing performance on those dimensions measured to the detriment of the dimensions that go unmeasured several times, he can be fairly sure that the principal's ToM do not allow to uncover the agent's behavior, and that it is possible to take shortcuts that maximize his payoffs. The agent's behavior over time thus is likely to revert back to the behavior predicted by standard agency models. In the short-term, however, principals might be tempted to signal a high ToM in order to stimulate the agent to work harder than when offered a flat salary. This may be particular attractive to principals who themselves are only agents of other principals, and who assume that they will move on to another position in a relatively short time frame. Under these circumstances, signaling a high ToM to the agent by choosing a $\beta > \beta^*$ allows these principals to reap the higher payoffs due to the agent working hard on all or most dimensions and exerting more effort than under a flat salary scheme. By the time the agent is reasonably

sure that the principal's true ToM abilities are lower than the ones necessary to allow the principal using such high-powered incentives, the principal may already have moved to another position in the firm or another firm. The higher-than-with-flat-salary performance of the agent on all dimensions of the task due to the $\beta > \beta^*$ might even increase the likelihood of the principal being promoted. Likewise, even from a firm perspective, choosing high incentive intensity for multi-tasking jobs may be reasonable if the firm relies on a strong job rotation practice among principals or agents. In this case, agents are unable to learn about a principal's ToM abilities since the principal for which they work regularly changes.

Taking account of the (imperfect) ToM of both principals and agents also opens up the door to a better understanding of a type of problem that in standard agency is largely disregarded: coordination problems (Foss, 2001). As Jacobides and Croson (2001: 211) note, “even agents who fully share the principal's objectives may dissipate joint agency value through miscoordination”. Imperfect ToM may entail such coordination problems. Agents may misread what the principal wants them to do and thus act post-contractually in a manner undesired by the principal. This danger may be particularly large if principal and agent are at different geographic locations and cannot meet face-to-face. In these cases, agents may not be able to collect the necessary insights – such as emotion expressions – to correctly read the principal's mind. Considering imperfect ToM thus opens up seeing the common lament among business practitioners about problems with strategy implementation not only in the case of incentive problems, but also with a broader view that allows for motivated agents to act in ways the principal did not want them to act. Bringing ToM into agency theory thus brings back the notion of effectiveness into the analysis of principal-agent relations – a notion so far largely overshadowed by a focus on minimizing inefficiency-related agency costs (see Pepper and Gore, 2013, for a discussion of the need to bring back effectiveness considerations into agency theory).

As these examples demonstrate, the insights gained when studying principal-agent relations with an agency theory that accounts for (imperfect) ToM of principals and agents differ from the standard model and offer new explanations for several important empirical phenomena.

IV. Conclusion

Numerous authors have criticized the motivational assumptions underlying the standard agency model. In contrast, the epistemic and cognitive assumptions have received far less attention and sparked little critical debate. However, as already the basic agency model shows, conflicting interests among principals and agents only become problematic if the parties involved in a trade differ in the information available to them. Therefore, the epistemic assumptions of agency theory are not less important to consider than the motivational ones when aiming at describing “average” principal-agent relationships.

The theory’s epistemic assumptions are highly asymmetrical (Foss and Hallberg, 2014): Whereas the principal is assumed to be perfectly knowledgeable of some pieces of information, others are considered to be perfectly private information of the agent. The theory thus largely models principal-agent relations in a binary manner. This binary view, however, is unlikely to correspond to the overwhelming majority of principal-agent relations in practice. As we have shown, most principals will have limited mentalizing capacities, implying that they may not have perfect knowledge of the agent’s risk aversion or costs of exerting effort. Therefore, finding the optimal incentive intensity in practice is a daunting task for most principals and chances are good that principals will err. Due to its epistemic and cognitive assumptions, standard agency theory assumes that the principal holds a perfect ToM with respect to the agent’s risk aversion and costs of exerting effort. For this reason, it risks underestimating agency costs.

Most real-life solutions to principal-agent problems therefore are likely better qualified as third-best rather than second-best. At the same time, however, standard agency theory

assumes fully imperfect ToM with respect to the agent's intentions as principals are assumed to be completely unable to ascertain which part of the output is due to the agent's efforts and which one is due to environmental factors beyond the agent's control. Most real-life principals, however, while unable (or unwilling due to high costs) to observe an agent's efforts, are likely to be able to put themselves to a certain degree into the agent's shoes. Interpreting the agent's body language or similar (non-verbal or verbal) cues overlooked in standard agency theory may allow them to at least somewhat decipher the true effort level of the agent. This implies that they incur lower agency costs than assumed in standard agency theory as it allows *ceteris paribus* to reduce the agent's exposure to risk – and, hence, to pay a lower risk premium – or, alternatively, to save on monitoring efforts.

Just like the recent literature approach on a “behavioral agency theory” aims at instilling the theory with more realistic assumptions about human motivation, risk preferences, and cognition (e.g., Pepper and Gore, 2013), incorporating the principal's and agent's ToM capacities into agency theory promises to lead to significantly refined conclusions and recommendations. We submit that this will help advance agency theory and further our understanding of real-life principal-relations.

References

- Adams, J. (1996). “Principals and Agents, Colonialists and Company Men: The Decay of Colonial Control in the Dutch East Indies.” *American Sociological Review* 61(February): 12-28.
- Ames, D. R., Weber, E. U., and Zou, X. (2012). “Mind-reading in strategic interaction: The impact of perceived similarity on projection and stereotyping.” *Organizational Behavior and Human Decision Processes* 117: 96-110.
- Aumann, R. J, and Brandenburger, A. (1995). “Epistemic conditions for Nash equilibrium.” *Econometrica* 63: 1161-1180.

- Baker, G., Gibbons, R., and Murphy, K. J. (1994). "Subjective performance measures in optimal incentive contracts." *Quarterly Journal of Economics*, 109: 1125-1156.
- Bagozzi, R. P., Verbeke, W. J. M. I., Berg, W. E., Rietdijk, W. J. R., Dietvorst, R. C., and Worm, L. (2013). "Genetic and neurological foundations of customer orientation: Field and experimental evidence." Forthcoming in: *Journal of the Academy of Marketing Science*.
- Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). "Does the autistic child have a 'theory of mind'?" *Cognition* 21: 37-46.
- Bewley, T. F. (1999). *Why Wages Don't Fall During a Recession*. Harvard University Press, Cambridge, MA.
- Camerer, C. (1998). "Bounded Rationality in Individual Decision Making." *Experimental Economics* 1: 163-183.
- Chng, D. H. M., Rodgers, M. S., Shih, E., and Song, X-B. (2012). "When does incentive compensation motivate managerial behaviors? An experimental investigation of the fit between incentive compensation, executive core self-evaluations, and firm performance." *Strategic Management Journal* 33(12): 1343-1362.
- Deci, E. L. (1971). "Effects of externally mediated rewards on intrinsic motivation." *Journal of Personality and Social Psychology* 18: 105-115.
- Deci, E. L., Koestner, R., and Ryan, R. M. (1999). "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation." *Psychological Bulletin* 125: 627-668.
- De Melo, C. M., Carnevale, P. J., Read, S. J., and Gratch, J. (2014). "Reading People's Minds from Emotion Expressions in Interdependent Decision Making." *Journal of Personality and Social Psychology* 106(1): 73-88.
- Dietvorst, R. C., Verbeke, W. J., Bagozzi, R. P., Yoon, C., Smits, M., and van der Lugt, A. (2009). "A sales force-specific theory-of-mind scale: Tests of its validity by classical

- methods and functional magnetic resonance imaging.” *Journal of Marketing Research*, 46: 653-668.
- Eisenhardt, K. M. (1989). “Agency theory: An assessment and review.” *Academy of Management Review* 14: 57-74.
- Ferraro, F., Pfeffer, J., and Sutton, R. I. (2005). “Economic Language and Assumptions: How Theories can become Self-fulfilling.” *Academy of Management Review* 30(1): 8-24.
- Fessler, N. J. (2003). “Experimental evidence on the links among monetary incentives, task attractiveness, and task performance.” *Journal of Management Accounting Research* 15: 161-176.
- Foss, N. J. (2001). “Leadership, Beliefs and Coordination: An Explorative Discussion.” *Industrial and Corporate Change* 10(2): 357-388.
- Foss, N. J. and N. Hallberg (2014). “How symmetrical assumptions facilitate theoretical advance in strategic management: The case of the resource-based view”. Forthcoming in: *Strategic Management Journal*.
- Foss, N. J., and Klein, P. G. (2012). *Organizing Entrepreneurial Judgment: A New Approach to the Firm*. Cambridge: Cambridge University Press.
- Foss, N. J., and Stea, D. (2014). “Putting a Realistic Theory of Mind into Agency Theory: Implications for Reward Design and Management in Principal-Agent Relations.” Forthcoming in *European Management Review*.
- Frith, U., and Frith, C. D. (2003). “Development and neurophysiology of mentalizing.” *Philosophical Transactions of the Royal Society of London, B: Biological Sciences* 358: 459-473.
- Gallagher, H. L., and Frith, C. D. (2003). “Functional imaging of ‘theory of mind’.” *Trends in Cognitive Sciences* 7: 77-83.
- Grossman, S. J., and Hart, O. D. (1983). “An analysis of the principal-agent problem.” *Econometrica*, 51: 7-45.

- Hein, G., and Singer, T. (2008). "I feel how you feel but not always: the empathic brain and its modulation." *Current Opinion on Neurobiology* 18(2): 153-158.
- Heath, C. (1999). "On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives." *Organizational Behavior and Human Decision Processes* 78: 25–62.
- Hendry, J. (2002). "The Principal's Other Problems: Honest Incompetence and the Specification of Objectives." *Academy of Management Review* 27: 98-113.
- Holler, M. J. (2002). "Classical, modern, and new game theory." [*Journal of Economics and Statistics* 222\(5\): 556-583.](#)
- Holmström, B. (1979). "Moral hazard and observability." *Bell Journal of Economics* 10: 74-91.
- Holmström, B. (1989). "Agency Costs and Innovation." *Journal of Economic Behavior and Organization* 12: 305-327.
- Holmström, B., and Milgrom, P. (1991). "Multitask principal agent analyses: Incentive contracts, asset ownership and job design." *Journal of Law, Economics and Organization* 7: 24-52.
- Jensen, M. C., and Meckling, W. H. (1976). "Theory of the firm: Managerial behavior, agency costs and ownership structure." *Journal of Financial Economics* 3: 305-360.
- Kahneman, D., and Tversky, A. (2000). *Choices, values and frames*. Cambridge: Cambridge University Press.
- Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Boston: Houghton Mifflin Company.
- Kunz, J., and Linder, S. (2012). "Organizational Control and Work Effort—Another Look at the Interplay of Rewards and Motivation." *European Accounting Review* 21(3): 591-621.
- Lazear, E. P. (2000). "Performance Pay and Productivity." *American Economic Review* 90: 1346-1361.

- Laffont, J.-J., and Martimort, D. (2001). *Theory of Incentives I: The Principal-Agent Model*. Princeton: Princeton University Press.
- Lepper, M. R., Greene, D., and Nisbett, R. E. (1973). "Undermining children's intrinsic interest with extrinsic reward: a test of the 'overjustification' hypothesis." *Journal of Personality and Social Psychology* 28(1): 129-137.
- Milgrom, P., and Roberts, J. (1992). *Economics, Organization & Management*. Englewood Cliffs: Prentice Hall.
- Pfeffer, J., and Langton, N. (1993). "The Effect of Wage Dispersion on Satisfaction, Productivity, and Working Collaboratively: Evidence from College and University Faculty." *Administrative Science Quarterly* 38(3): 382-407.
- Pepper, A., and Gore J. (2013). "The Economic Psychology of Incentives: An international Study of Top Managers." Forthcoming in: *Journal of World Business*.
- Perrow, C. (1990). "Economic Theories of Organization." In: Zukin S, Dimaggio P (eds.) *Structures of Capital*. Cambridge: Cambridge University Press.
- Premack, D., and Woodruff, G. (1978). "Does the chimpanzee have a theory of mind?" *Behavioral and Brain Sciences* 1: 515-526.
- Prendergast, C., and Topel, R. H. (1996). "Favoritism in Organizations." *Journal of Political Economy* 104(5): 958-978.
- Robalino, N., and Robson, A. (2012). "The economic approach to 'theory of mind'." *Philosophical Transactions of The Royal Society - B*, 367: 2224-2233.
- Rose-Ackerman, S. (1975). "The Economics of Corruption." *Journal of Public Economics* 4(2): 187-203.
- Ross, S. A. (1973). "The Economic Theory of Agency: The Principal's Problem." *American Economic Review* 63(2): 134-39.
- Singer, T., and Fehr, E. (2005). "The neuroeconomics of mind reading and empathy." *American Economic Review* 95: 340-345.

Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks: Sage.

Wowak, A. J., and Hambrick, D. C. (2010). "A model of person-pay interaction: how executives vary in their responses to compensation arrangements." *Strategic Management Journal* 31(8): 803-821.

Table 1: What standard agency theory assumes principal and agent to know

| Piece of Information | Principal | Agent | Remark |
|---|-----------|-------|---|
| agent's skills and abilities (set, quality, etc.) | ✗ | ✓ | "hidden characteristics" |
| agent's space of actions/effort e | ✓ | ✓ | complete contracts |
| actual e | ✗ | ✓ | "hidden action" |
| $C''(e)$ | ✓ | ✓ | |
| r_A | ✓ | ✓ | r_A is independent of wealth |
| $Exp(\theta)$ | ✓ | ✓ | |
| Distribution and $Var(\theta)$ | ✓ | ✓ | |
| actual θ | ✗ | ✓ | |
| production function | ✓ | ✓ | linear, additive function $z = e + \theta$, with agent having no influence over θ , assumed in basic models |
| $W = W_0 + \beta z$ | ✓ | ✓ | |
| $Exp[Utility(W, e)] \geq \underline{U}$ for all e | ✓ | ✓ | "Participation constraint" |
| self-interest of the agent | ✓ | ✓ | assumed to be "bad" |
| self-interest of principal | ✓ | | assumed to be "good" |
| Object the conflicting interests pertain to (e.g., outcome type, time preference, etc.) | ✓ | ✓ | leisure vs. effort in the basic models |
| Surplus available through trade | ✓ | ✓ | both parties recognize potential gains from trade |
| Rationality of agent | ✓ | ✓ | |
| Rationality of principal | ✓ | ✓ | |
| Rationality of respective other party | ✓ | ✓ | |
| Sensitivity of principal's payoffs to change in e ($P'(e)$) | ✓ | | |
| Principal's own risk aversion | ✓ | | zero in the basic models |
| Principal's costs of monitoring | ✓ | | zero in the basic models |

Note: ✓ denotes that the standard model assumes the player to know (✗ not to know) this piece of information; a blank cell indicates that the standard model does not take an explicit stance; $P'(e)$ and $C''(e)$ denote the first/respectively the second derivatives of $P(e)$ and $C(e)$ to e respectively.

SMG – Working Papers

www.cbs.dk/smg

2003

- 2003-1:** Nicolai J. Foss, Kenneth Husted, Snejina Michailova, and Torben Pedersen: Governing Knowledge Processes: Theoretical Foundations and Research Opportunities.
- 2003-2:** Yves Doz, Nicolai J. Foss, Stefanie Lenway, Marjorie Lyles, Silvia Massini, Thomas P. Murtha and Torben Pedersen: Future Frontiers in International Management Research: Innovation, Knowledge Creation, and Change in Multinational Companies.
- 2003-3:** Snejina Michailova and Kate Hutchings: The Impact of In-Groups and Out-Groups on Knowledge Sharing in Russia and China CKG Working Paper.
- 2003-4:** Nicolai J. Foss and Torben Pedersen: The MNC as a Knowledge Structure: The Roles of Knowledge Sources and Organizational Instruments in MNC Knowledge Management CKG Working Paper.
- 2003-5:** Kirsten Foss, Nicolai J. Foss and Xosé H. Vázquez-Vicente: “Tying the Manager’s Hands”: How Firms Can Make Credible Commitments That Make Opportunistic Managerial Intervention Less Likely CKG Working Paper.
- 2003-6:** Marjorie Lyles, Torben Pedersen and Bent Petersen: Knowledge Gaps: The Case of Knowledge about Foreign Entry.
- 2003-7:** Kirsten Foss and Nicolai J. Foss: The Limits to Designed Orders: Authority under “Distributed Knowledge” CKG Working Paper.
- 2003-8:** Jens Gammelgaard and Torben Pedersen: Internal versus External Knowledge Sourcing of Subsidiaries - An Organizational Trade-Off.
- 2003-9:** Kate Hutchings and Snejina Michailova: Facilitating Knowledge Sharing in Russian and Chinese Subsidiaries: The Importance of Groups and Personal Networks Accepted for publication in *Journal of Knowledge Management*.
- 2003-10:** Volker Mahnke, Torben Pedersen and Markus Verzin: The Impact of Knowledge Management on MNC Subsidiary Performance: the Role of Absorptive Capacity CKG Working Paper.
- 2003-11:** Tomas Hellström and Kenneth Husted: Mapping Knowledge and Intellectual Capital in Academic Environments: A Focus Group Study Accepted for publication in *Journal of Intellectual Capital* CKG Working Paper.
- 2003-12:** Nicolai J Foss: Cognition and Motivation in the Theory of the Firm: Interaction or “Never the Twain Shall Meet”? Accepted for publication in *Journal des Economistes et des Etudes Humaines* CKG Working Paper.
- 2003-13:** Dana Minbaeva and Snejina Michailova: Knowledge Transfer and Expatriation Practices in MNCs: The Role of Disseminative Capacity.
- 2003-14:** Christian Vintergaard and Kenneth Husted: Enhancing Selective Capacity Through Venture Bases.

2004

- 2004-1:** Nicolai J. Foss: Knowledge and Organization in the Theory of the Multinational Corporation: Some Foundational Issues
- 2004-2:** Dana B. Minbaeva: HRM Practices and MNC Knowledge Transfer
- 2004-3:** Bo Bernhard Nielsen and Snejina Michailova: Toward a Phase-Model of Global Knowledge Management Systems in Multinational Corporations
- 2004-4:** Kirsten Foss & Nicolai J Foss: The Next Step in the Evolution of the RBV: Integration with Transaction Cost Economics
- 2004-5:** Teppo Felin & Nicolai J. Foss: Methodological Individualism and the Organizational Capabilities Approach
- 2004-6:** Jens Gammelgaard, Kenneth Husted, Snejina Michailova: Knowledge-sharing Behavior and Post-acquisition Integration Failure
- 2004-7:** Jens Gammelgaard: Multinational Exploration of Acquired R&D Activities
- 2004-8:** Christoph Dörrenbächer & Jens Gammelgaard: Subsidiary Upgrading? Strategic Inertia in the Development of German-owned Subsidiaries in Hungary
- 2004-9:** Kirsten Foss & Nicolai J. Foss: Resources and Transaction Costs: How the Economics of Property Rights Furthers the Resource-based View
- 2004-10:** Jens Gammelgaard & Thomas Ritter: The Knowledge Retrieval Matrix: Codification and Personification as Separate Strategies
- 2004-11:** Nicolai J. Foss & Peter G. Klein: Entrepreneurship and the Economic Theory of the Firm: Any Gains from Trade?
- 2004-12:** Akshey Gupta & Snejina Michailova: Knowledge Sharing in Knowledge-Intensive Firms: Opportunities and Limitations of Knowledge Codification
- 2004-13:** Snejina Michailova & Kate Hutchings: Knowledge Sharing and National Culture: A Comparison Between China and Russia

2005

- 2005-1:** Keld Laursen & Ammon Salter: My Precious - The Role of Appropriability Strategies in Shaping Innovative Performance
- 2005-2:** Nicolai J. Foss & Peter G. Klein: The Theory of the Firm and Its Critics: A Stocktaking and Assessment
- 2005-3:** Lars Bo Jeppesen & Lars Frederiksen: Why Firm-Established User Communities Work for Innovation: The Personal Attributes of Innovative Users in the Case of Computer-Controlled Music
- 2005-4:** Dana B. Minbaeva: Negative Impact of HRM Complementarity on Knowledge Transfer in MNCs
- 2005-5:** Kirsten Foss, Nicolai J. Foss, Peter G. Klein & Sandra K. Klein: Austrian Capital

Theory and the Link Between Entrepreneurship and the Theory of the Firm

- 2005-1:** Nicolai J. Foss: The Knowledge Governance Approach
- 2005-2:** Torben J. Andersen: Capital Structure, Environmental Dynamism, Innovation Strategy, and Strategic Risk Management
- 2005-3:** Torben J. Andersen: A Strategic Risk Management Framework for Multinational Enterprise
- 2005-4:** Peter Holdt Christensen: Facilitating Knowledge Sharing: A Conceptual Framework
- 2005-5:** Kirsten Foss & Nicolai J. Foss: Hands Off! How Organizational Design Can Make Delegation Credible
- 2005-6:** Marjorie A. Lyles, Torben Pedersen & Bent Petersen: Closing the Knowledge Gap in Foreign Markets - A Learning Perspective
- 2005-7:** Christian Geisler Asmussen, Torben Pedersen & Bent Petersen: How do we Capture "Global Specialization" when Measuring Firms' Degree of internationalization?
- 2005-8:** Kirsten Foss & Nicolai J. Foss: Simon on Problem-Solving: Implications for New Organizational Forms
- 2005-9:** Birgitte Grøgaard, Carmine Gioia & Gabriel R.G. Benito: An Empirical Investigation of the Role of Industry Factors in the Internationalization Patterns of Firms
- 2005-10:** Torben J. Andersen: The Performance and Risk Management Implications of Multinationality: An Industry Perspective
- 2005-11:** Nicolai J. Foss: The Scientific Progress in Strategic Management: The case of the Resource-based view
- 2005-12:** Koen H. Heimeriks: Alliance Capability as a Mediator Between Experience and Alliance Performance: An Empirical Investigation Into the Alliance Capability Development Process
- 2005-13:** Koen H. Heimeriks, Geert Duysters & Wim Vanhaverbeke: Developing Alliance Capabilities: An Empirical Study
- 2005-14:** JC Spender: Management, Rational or Creative? A Knowledge-Based Discussion

2006

- 2006-1:** Nicolai J. Foss & Peter G. Klein: The Emergence of the Modern Theory of the Firm
- 2006-2:** Teppo Felin & Nicolai J. Foss: Individuals and Organizations: Thoughts on a Micro-Foundations Project for Strategic Management and Organizational Analysis
- 2006-3:** Volker Mahnke, Torben Pedersen & Markus Venzin: Does Knowledge Sharing

Pay? An MNC Subsidiary Perspective on Knowledge Outflows

- 2006-4:** Torben Pedersen: Determining Factors of Subsidiary Development
- 2006-5** Ibuki Ishikawa: The Source of Competitive Advantage and Entrepreneurial Judgment in the RBV: Insights from the Austrian School Perspective
- 2006-6** Nicolai J. Foss & Ibuki Ishikawa: Towards a Dynamic Resource-Based View: Insights from Austrian Capital and Entrepreneurship Theory
- 2006-7** Kirsten Foss & Nicolai J. Foss: Entrepreneurship, Transaction Costs, and Resource Attributes
- 2006-8** Kirsten Foss, Nicolai J. Foss & Peter G. Klein: Original and Derived Judgement: An Entrepreneurial Theory of Economic Organization
- 2006-9** Mia Reinholt: No More Polarization, Please! Towards a More Nuanced Perspective on Motivation in Organizations
- 2006-10** Angelika Lindstrand, Sara Melen & Emilia Rovira: Turning social capital into business? A study of Swedish biotech firms' international expansion
- 2006-11** Christian Geisler Asmussen, Torben Pedersen & Charles Dhanaraj: Evolution of Subsidiary Competences: Extending the Diamond Network Model
- 2006-12** John Holt, William R. Purcell, Sidney J. Gray & Torben Pedersen: Decision Factors Influencing MNEs Regional Headquarters Location Selection Strategies
- 2006-13** Peter Maskell, Torben Pedersen, Bent Petersen & Jens Dick-Nielsen: Learning Paths to Offshore Outsourcing - From Cost Reduction to Knowledge Seeking
- 2006-14** Christian Geisler Asmussen: Local, Regional or Global? Quantifying MNC Geographic Scope
- 2006-15** Christian Bjørnskov & Nicolai J. Foss: Economic Freedom and Entrepreneurial Activity: Some Cross-Country Evidence
- 2006-16** Nicolai J. Foss & Giampaolo Garzarelli: Institutions as Knowledge Capital: Ludwig M. Lachmann's Interpretative Institutionalism
- 2006-17** Koen H. Heimriks & Jeffrey J. Reuer: How to Build Alliance Capabilities
- 2006-18** Nicolai J. Foss, Peter G. Klein, Yasemin Y. Kor & Joseph T. Mahoney: Entrepreneurship, Subjectivism, and the Resource - Based View: Towards a New Synthesis
- 2006-19** Steven Globerman & Bo B. Nielsen: Equity Versus Non-Equity International Strategic Alliances: The Role of Host Country Governance

2007

- 2007-1** Peter Abell, Teppo Felin & Nicolai J. Foss: Building Micro-Foundations for the Routines, Capabilities, and Performance Links

- 2007-2** Michael W. Hansen, Torben Pedersen & Bent Petersen: MNC Strategies and Linkage Effects in Developing Countries
- 2007-3** Niron Hashai, Christian G. Asmussen, Gabriel R.G. Benito & Bent Petersen: Predicting the Diversity of Foreign Entry Modes
- 2007-4** Peter D. Ørberg Jensen & Torben Pedersen: Whether and What to Offshore?
- 2007-5** Ram Mudambi & Torben Pedersen: Agency Theory and Resource Dependency Theory: Complementary Explanations for Subsidiary Power in Multinational Corporations
- 2007-6** Nicolai J. Foss: Strategic Belief Management
- 2007-7** Nicolai J. Foss: Theory of Science Perspectives on Strategic Management Research: Debates and a Novel View
- 2007-8** Dana B. Minbaeva: HRM Practices and Knowledge Transfer in MNCs
- 2007-9** Nicolai J. Foss: Knowledge Governance in a Dynamic Global Context: The Center for Strategic Management and Globalization at the Copenhagen Business School
- 2007-10** Paola Gritti & Nicolai J. Foss: Customer Satisfaction and Competencies: An Econometric Study of an Italian Bank
- 2007-11** Nicolai J. Foss & Peter G. Klein: Organizational Governance
- 2007-12** Torben Juul Andersen & Bo Bernhard Nielsen: The Effective Ambidextrous Organization: A Model of Integrative Strategy Making Processes.

2008

- 2008-1** Kirsten Foss & Nicolai J. Foss: Managerial Authority When Knowledge is Distributed: A Knowledge Governance Perspective
- 2008-2** Nicolai J. Foss: Human Capital and Transaction Cost Economics.
- 2008-3** Nicolai J. Foss & Peter G. Klein: Entrepreneurship and Heterogeneous Capital.
- 2008-4** Nicolai J. Foss & Peter G. Klein: The Need for an Entrepreneurial Theory of the Firm.
- 2008-5** Nicolai J. Foss & Peter G. Klein: Entrepreneurship: From Opportunity Discovery to Judgment.
- 2008-6** Mie Harder: How do Rewards and Management Styles Influence the Motivation to Share Knowledge?
- 2008-7** Bent Petersen, Lawrence S. Welch & Gabriel R.G. Benito: Managing the Internalisation Process – A Theoretical Perspective.
- 2008-8** Torben Juul Andersen: Multinational Performance and Risk Management Effects: Capital Structure Contingencies.

- 2008-9** Bo Bernard Nielsen: Strategic Fit and the Role of Contractual and Procedural Governance in Alliances: A Dynamic Perspective.
- 2008-10** Line Gry Knudsen & Bo Bernhard Nielsen: Collaborative Capability in R&D Alliances: Exploring the Link between Organizational and Individual level Factors.
- 2008-11** Torben Juul Andersen & Mahesh P. Joshi: Strategic Orientations of Internationalizing Firms: A Comparative Analysis of Firms Operating in Technology Intensive and Common Goods Industries.
- 2008-12** Dana Minbaeva: HRM Practices Affecting Extrinsic and Intrinsic Motivation of Knowledge Receivers and their Effect on Intra-MNC Knowledge Transfer.
- 2008-13** Steen E. Navrbjerg & Dana Minbaeva: HRM and IR in Multinational Corporations: Uneasy Bedfellows?
- 2008-14** Kirsten Foss & Nicolai J. Foss: Hayekian Knowledge Problems in Organizational Theory.
- 2008-15** Torben Juul Andersen: Multinational Performance Relationships and Industry Context.
- 2008-16** Larissa Rabbiosi: The Impact of Subsidiary Autonomy on MNE Knowledge Transfer: Resolving the Debate.
- 2008-17** Line Gry Knudsen & Bo Bernhard Nielsen: Organizational and Individual Level Antecedents of Procedural Governance in Knowledge Sharing Alliances.
- 2008-18** Kirsten Foss & Nicolai J. Foss: Understanding Opportunity Discovery and Sustainable Advantage: The Role of Transaction Costs and Property Rights.
- 2008-19** Teppo Felin & Nicolai J. Foss: Social Reality, The Boundaries of Self-fulfilling Prophecy, and Economics.
- 2008-20** Yves Dos, Nicolai J. Foss & José Santos: A Knowledge System Approach to the Multinational Company: Conceptual Grounding and Implications for Research
- 2008-21** Sabina Nielsen & Bo Bernhard Nielsen: Why do Firms Employ foreigners on Their Top Management Teams? A Multi-Level Exploration of Individual and Firm Level Antecedents
- 2008-22** Nicolai J. Foss: Review of Anders Christian Hansen's "Uden for hovedstrømmen – Alternative strømninger i økonomisk teori"
- 2008-23** Nicolai J. Foss: Knowledge, Economic Organization, and Property Rights
- 2008-24** Sjoerd Beugelsdijk, Torben Pedersen & Bent Petersen: Is There a Trend Towards Global Value Chain Specialization? – An Examination of Cross Border Sales of US Foreign Affiliates

- 2008-25** Vikas Kumar, Torben Pedersen & Alessandro Zattoni: The performance of business group firms during institutional transition: A longitudinal study of Indian firms
- 2008-26** Sabina Nielsen & Bo B. Nielsen: The effects of TMT and Board Nationality Diversity and Compensation on Firm Performance
- 2008-27** Bo B. Nielsen & Sabina Nielsen: International Diversification Strategy and Firm Performance: A Multi-Level Analysis of Firm and Home Country Effects

2009

- 2009-1** Nicolai J. Foss: Alternative Research Strategies in the Knowledge Movement: From Macro Bias to Micro-Foundations and Multi-Level Explanation
- 2009-2** Nicolai J. Foss & Peter G. Klein: Entrepreneurial Alertness and Opportunity Discovery: Origins, Attributes, Critique
- 2009-3** Nicolai J. Foss & Dana B. Minbaeva: Governing Knowledge: The Strategic Human Resource Management Dimension
- 2009-4** Nils Stieglitz & Nicolai J. Foss: Opportunities and New Business Models: Transaction Cost and Property Rights Perspectives on Entrepreneurships
- 2009-5** Torben Pedersen: Vestas Wind Systems A/S: Exploiting Global R&D Synergies
- 2009-6** Rajshree Agarwal, Jay B. Barney, Nicolai J. Foss & Peter G. Klein: Heterogeneous Resources and the Financial Crisis: Implications of Strategic Management Theory
- 2009-7** Jasper J. Hotho: A Measure of Comparative Institutional Distance
- 2009-8** Bo B. Nielsen & Sabina Nielsen: The Impact of Top Management Team Nationality Diversity and International Experience on Foreign Entry Mode
- 2009-9** Teppo Felin & Nicolai Juul Foss: Experience and Repetition as Antecedents of Organizational Routines and Capabilities: A Critique of Behaviorist and Empiricist Approaches
- 2009-10** Henk W. Volberda, Nicolai J. Foss & Marjorie E. Lyles: Absorbing the Concept of Absorptive Capacity: How To Realize Its Potential in the Organization Field
- 2009-11** Jan Stentoft Arlbjørn, Brian Vejrum Wæhrens, John Johansen & Torben Pedersen: Produktion i Danmark eller offshoring/outsourcing: Ledelsesmæssige udfordringer
- 2009-12** Torben Pedersen: The 30 Largest Firms in Denmark

2010

- 2010-1** Dana B. Minbaeva, Kristiina Mäkelä & Larissa Rabbiosi: Explaining Intra-organizational Knowledge Transfer at the Individual Level
- 2010-2** Dana B.Minbaeva & Torben Pedersen: Governing Individual Knowledge Sharing Behavior
- 2010-3** Nicolai J. Foss & Peter G. Klein: Alertness, Judgment, and the Antecedents of Entrepreneurship
- 2010-4** Nicolai J.Foss & Joseph T.Mahoney: Exploring Knowledge Governance
- 2010-5** Jasper J. Hotho, Florian Becker-Ritterspach & Ayse Saka-Helmhout: Enriching Absorptive Capacity Through Social Interaction
- 2010-6** Nicolai J. Foss & Bo B. Nielsen: Researching Collaborative Advantage: Some Conceptual and Multi-level Issues
- 2010-7** Nicolai J. Foss & Nils Stieglitz: Modern Resource-Based Theory(ies)
- 2010-8** Christian Bjørnskov & Nicolai J. Foss: Do Economic Freedom and Entrepreneurship Impact Total Factor Productivity?
- 2010-9** Gabriel R.G. Benito, Bent Petersen & Lawrence S. Welch: Mode Combinations and International Operations: Theoretical Issues and an Empirical Investigation

2011

- 2011-1** Peter D. Ørberg Jensen & Bent Petersen: Human Asset Internalization and Global Sourcing of Services – A Strategic Management Analysis on Activity-level
- 2011-2** Mie Harder: Management Innovation Capabilities: A Typology and Propositions for Management Innovation Research
- 2011-3** Mie Harder: Internal Antecedents of Management Innovation: The effect of diagnostic capability and implementation capability

- 2011-4** Mie Harder: Explaining Management Innovation Pervasiveness: The Role of Internal Antecedents
- 2011-5** Mie Harder: Internal Determinants of Product Innovation and Management Innovation: The Effect of Diagnostic Capability and Implementation Capability
- 2011-6** Nicolai J. Foss, Peter G. Klein & Per L. Bylund: Entrepreneurship and the Economics of the Firm
- 2011-7** Nicolai J. Foss & Jacob Lyngsie: The Emerging Strategic Entrepreneurship Field: Origins, Key Tenets and Research Gaps
- 2011-8** Nicolai J. Foss: Entrepreneurship in the Context of the Resource-based View of the Firm
- 2011-9** Bent Petersen, Gabriel R.G. Benito, Olesya Dovgan & Lawrence Welch: Offshore outsourcing: A dynamic, operation mode perspective
- 2011-10** Bent Petersen, Gabriel R. G. Benito & Lawrence Welch: Dynamics of Foreign Operation Modes and their Combinations: Insights for International Strategic Management
- 2011-11** Nicolai J. Foss: Teams, Team Motivation, and the Theory of the Firm
- 2011-12** Nicolai J. Foss: Knowledge Governance: Meaning, Nature, Origins, and Implications
- 2011-13** Nicolai J. Foss, Kirsten Foss & Phillip C. Nell: MNC Organizational Form and Subsidiary Motivation Problems: Controlling Intervention Hazards in the Network MNC
- 2011-14** Kåre Moberg: Evaluating Content Dimensions in Entrepreneurship Education

2012

- 2012-1** Nicolai J. Foss, Nicholas Argyres, Teppo Felin & Todd Zenger: The Organizational Economics of Organizational Capability and Heterogeneity: A Research Agenda

- 2012-2** Torben J. Andersen, Carina Antonia Hallin & Sigbjørn Tveterås: A Prediction Contest: The Sensing of Frontline Employees Against Executive Expectations
- 2012-3** Peter G. Klein, Jay B. Barney & Nicolai J. Foss: Strategic Entrepreneurship
- 2012-4** Kåre Moberg: The Impact of Entrepreneurship Education and Project-based Education on Students' Personal Development and Entrepreneurial Intentions at the Lower Levels of the Educational System: Too Much of Two Good Things?
- 2012-5** Keld Laursen & Nicolai J. Foss: Human Resource Management Practices and Innovation
- 2012-6** Kåre Moberg: An Entrepreneurial Self-Efficacy Scale with a Neutral Wording

2013

- 2013-1**
Nicolai J. Foss, Diego Stea: The Principal's Theory of Mind: The Role of Mentalizing for Reward Design and Management in Principal-Agent Relations
- 2013-2** Dana Minbaeva, Chansoo Park & Ilan Vertinsky: The Influence of Foreign Partners' Disseminative Capacities on Knowledge Transfers to International Joint Ventures
- 2013-3** Nicolai J. Foss & Peter G. Klein: Hayek and Organizational Studies
- 2013-4** Kåre Moberg, Lene Vestergaard, Casper Jørgensen, Elisabeth Markussen & Sose Hakverdyan: How to Assess the Development of Entrepreneurship Education at University Level – the Case of Denmark
- 2013-5** Nicolai J. Foss & Siegwart Lindenberg: Micro-Foundations For Strategy: A Goal-Framing Perspective on the Drivers of Value Creation

- 2013-6 Nicolai J. Foss, Jacob Lyngsie & Shaker A. Zahra: The Role of External Knowledge Sources and Organizational Design in the Process of Opportunity Exploitation
- 2013-7 Stefan Linder & Nicolai J. Foss: Agency Theory.
- 2013-8 Nicolai J. Foss, Peter G. Klein, Stefan Linder: Organizations and Markets
- 2013-9 Nicolai J. Foss: Towards an Organizational Economics of Heterogeneous Capabilities
- 2013- 10 Christian Geisler Asmussen & Nicolai J. Foss: Competitive Advantage and the Existence of the MNC: Earlier Research and the Role of Frictions

2014

- 2014-1 Nicolai J. Foss & Kirsten Foss: COASIAN AND MODERN PROPERTY RIGHTS ECONOMICS: A CASE OF KUHNIAN LOST CONTENT
- 2014-2 Nicolai J. Foss & Nils Stieglitz: BUSINESS MODEL INNOVATION: THE ROLE OF LEADERSHIP
- 2014-3 Nicolai J. Foss, Stefan Linder & Diego Stea: EPISTEMICS AT WORK: THE THEORY OF MIND IN PRINCIPAL-AGENT RELATIONS