Transaction Cost Economics and Plausible Actors: A Cognitive Reappraisal[1]

Luit Gazendam & René J. Jorna
Faculty of Management & Organization
University of Groningen, Groningen, The Netherlands
E-mail: r.j.j.m.jorna@bdk.rug.nl

Abstract
This article discusses the cognitive plausibility of actors or agents within transaction cost economics (TCE). Our main thesis is that, although the actors within TCE are not based on cognitive functions, a simulation model can be built in which the cognition of the actors is accounted for. In this article we want to discuss four things. In the first place we describe TCE and its actors in which we especially analyze actor characteristics, such as trust, opportunism, bounded rationality, loyalty and learning. We conclude that these actor characteristics lack cognitive psychological plausibility. In the second place we give an overview of kinds of actors, which make it possible to discern the actor model TCE is using from other actor models. We make a distinction in actors as response function systems, as representational systems and as representational response function systems. In the third place we indicate that for a theory such as TCE in which complex actors and governance structures are studied, at least two levels of description can be determined: an intentional and an functional level. Within TCE trust, opportunism and loyalty of actors are described at an intentional level, without any reference to the underlying actor mechanisms which are in terms of functional descriptions. In order to complement TCE actors have to be made more cognitively plausible. In the fourth place we give an outline of an architecture for cognition - the ACT-R approach - that will be used to implement in an actor mechanisms of trust, loyalty, opportunism and learning. This article ends with a conceptual structure of a TCE-COG actor that can be used in a simulation to put TCE to a better test. The real implementation of ACT-R actors in a TCE simulation is part of ongoing research.

1. We would like to thank Martin Helmhout and Diederik Kraaijkamp for comments on earlier drafts.

Transaction Cost Economics and Plausible Actors: A Cognitive Reappraisal

1. Introduction: Transaction Cost Economics, Trust and Cognition

Economists try to model the (economic) behavior of people in firms and in organizations as best as possible. Models and theories are used that describe the dynamics of - or the equilibria in - the economic behavior of people and firms. A well known example of such an economic theory is transaction cost economics (TCE). In TCE the unit of analysis is the transaction and the firm is considered to be one of the alternatives for organizing these transactions. With every organization of transactions, costs are involved that vary with the properties of the transaction. Because different organizational forms vary in the way they handle these costs, the one organizational form is more suited for organizing transactions with specific properties than the other.

In TCE the central focus is on transactions. A transaction is said to take place when a good or service passes a technological barrier. The barrier determines where one phase of activity ends and another begins. TCE does not focus on the optimization of individual phases of activity - where the firm is the production function which needs optimization -, but on the transactions between two phases of activity. However, TCE also is an optimization theory. With every transaction, costs are involved. Within the firm this may include coordination and organizational costs and outside the firm contracting costs and contract fulfilment costs between two firms. Profit is attainable by optimizing the expected revenues with respect to the costs to be made. It is assumed that firms try to optimize their profit, but they cannot be sure that they will succeed. Therefore, TCE also accounts for external influences on the market. In the struggle for the survival of the fittest on the market those organizational forms that are optimal for the organization of certain transactions will turn out to be the best. Even though firms try to optimize, the process of optimization is carried out by the market and not by the firms themselves.

We can imagine a transaction within TCE by taking an arrow that goes from A to B. We can ask ourselves what happens at the point of departure and at the point of arrival of this arrow. At these points TCE starts with individual actors. In TCE the transaction is always conceptualized as a transaction between two individual actors. This means that a transaction, whether between two firms or between two departments of one firm, in the end always boils down to individuals. Of course the individual actors can be part of bigger units that constitute the firm. It is also possible that an individual actor and a firm coincide.

Within TCE the focus on transactions between individual actors also requires some additional assumptions, that are incorporated in concepts. Some concepts apply to the organization of transactions, such as governance structures, others apply to the conceptualisation of transactions between the individual actors, such as trust, opportunism and bounded rationality. The latter concepts are related to the characteristics of individual actors. Although TCE focuses on the transaction, key concepts are related to the characteristics of the individual actors; they are the starting and ending points. In this actor perspective it is the question whether TCE really starts from psychological assumptions, and if so in what sense are these assumptions corroborated by empirical and theoretical evidence. If for a transaction to take place, an individual actor needs trust or has loyalty with respect to another actor with whom a transaction is made, an important question is how mechanisms for trust and loyalty (implicitly or explicitly) are conceptualized within TCE's assumptions for individual actors?

In this article we will focus on several characteristics of actors, such as trust, loyalty, opportunism, learning and bounded rationality. Other aspects that are related to governance and coordination will not be discussed in detail here. Influences on transactions by colleagues or the organization itself, that may complicate the analysis of transactions, will also be left out. Furthermore, we assume that an actor and a human individual coincide, for even if transactions take place between organizations, in the end always an individual actor is involved (Jorna 2000).

Because a theory like TCE, just as almost all economic theories that are based on methodological individualism, starts from individual actors, we want to focus on the psychological assumptions of TCE. With respect to these assumptions we will choose a cognitive perspective. Just as TCE we conceive actors to be intelligent, which means that we depart from actors as cognitive systems, or in other words, as information processing systems (Newell, 1990).

The structure of this article is as follows. In section 2 we will discuss TCE-actors with respect to opportunism, loyalty, learning, bounded rationality and trust. We will concentrate on the flaws in these assumptions. We will explain that TCE does not include mechanisms which allow for the appearance of trust or loyalty. This should not be considered as a flaw of economy, but as a consequence of the distribution of scientific labor. Cognitive scientists have to come up with plausible (cognitive) models of actors. Economists should continue to built upon these models in explaining economic behavior. In case economists use other models it should be clear what these actor models include and perhaps how they can be improved. A larger part of section 2 will be devoted to the meaning of trust. In section 3 we will make some general remarks about kinds of actors and how they can be arranged. This will give an underlying structure for a cognitive reappraisal of TCE-actors. In section 4 we will argue that cognitive science has a large set of possible cognitive actor models, from which we will choose the ACT-R model. We will explain why this choice is made and what the consequences are for TCE. The role of artificial intelligence will be discussed in section 5. In section 6 we will discuss ACT-R in detail and a possible implementation of ACT-R actors in the simulation of Klos (2000) and Helmhout (2001) and in section 7 we will give some conclusions.

2 Loyalty, trust, rationality, opportunism and learning within TCE

2.1 Introduction in TCE

In the configuration of a simulation model with several TCE actors, that behave like buyers and suppliers, Klos (2000) in chapter 1 summarizes key notions within TCE as follows:

"Transaction cost economics is often used to analyze the issue of make or buy, but many other questions of organising as well. This theory originated in the paper "The Nature of the Firm" by Coase (1937) and was subsequently developed by, among others, Williamson (1975, 1985). Coase (1937, p.390) wondered "why a firm emerges at all in a specialised exchange economy", and "why coordination is the work of the price mechanism in one case and of the entrepreneur in another" (1937, p. 389). According to Coase and transaction cost economics in general, the firm should not be studied in isolation as a production function to be optimized, but the transaction should be made the basic unit of analysis and the firm should be seen as one of several alternative ways to organize transactions. Costs are associated with transactions that vary with the characteristics of the transaction. Because organizational forms are differentially able to economize on those costs, different organizational forms are more suited for organizing transactions with particular characteristics than for others. Transaction cost economics is about finding the most appropriate – i.e. the economic – organizational form given the characteristics of a transaction. The relevant characteristics are the frequency of the transaction, the uncertainty surrounding it and the transaction specificity of the assets invested in it. In its search for optimal, economic organizational forms, TCE assumes that people are bounded rational as well as potentially opportunistic. This first assumption means that people are unable to gather as well as process all the information required to make optimal decisions. The second means that people do not simply seek self-interest, but that they do so with guile, including "calculated efforts to mislead, disguise, obfuscate, or otherwise confuse" (Williamson 1985, p. 47), and that

differential trustworthiness is hard, if not impossible to predict." (Klos 2000, Chapter 1)

Klos leaves out the details of what is said about actors, because he is not interested in the (cognitive) psychological determination of the actor. Instead he focuses on an simulation in which trust and loyalty of buyers and suppliers are accounted for from a non-cognitive perspective.

2.2 Individually bound aspects of TCE

The individually bound aspects of TCE refer to individualism and governance, loyalty and opportunism, (bounded) rationality and trust. In the next sub-sections we will discuss these aspects in more detail.

2.2.1 Individuality and governance

As said before, TCE claims that transactions can take place within the firm or between two firms. In both cases the transactions are established between two persons and are shaped by those persons: "*The logic of this argumentation hinges on its methodological individualism. In all cases economic interaction is conceptualized as transactions between individual actors*." (Noorderhaven, 1993, p. 7) "*The central place of methodological individualism is also reflected in the use of the concept of [unified] contracting. [...] How can the unified governance structure safeguard transactions if the parties remain relentlessly opportunistic*?" (Ibidem, p. 7-8). When a transaction takes places within a firm, people - regardless of their position within the hierarchy - are supposed to have a joint goal, i.e., to pass off the transaction as cheap and as smooth as possible. They will do this because people have the feeling that they are together in the firm. "*Williamson at last introduces two arguments that are to make plausible the superiority of unified governance in curbing opportunism: "atmosphere" and "consummate cooperation". Atmosphere (...) implies that organizational modes may not only differ in efficiency respects, but may also engender differences in attitudes. ...Loose allusions are made to the "sense that managers and workers are in this together" and the benefits of "employee loyalty" can be associated with "atmosphere", but this connection is not made explicit.*" (Ibidem, p. 11) "*The spirit in which adaptations are effected is [...] important.*" (Ibidem, p. 16) This means that integrity and attitude have an important influence on the behavior of people within a firm. Furthermore, it is argued within TCE that when a transaction takes place within the firm, people always cooperate. However, when a transaction takes place between two firms, people - and thus firms - are assumed to employ opportunistic behavior. With regard to opportunism, within TCE it is assumed that "*opportunism forms the invariant core of economic actors; trustworthiness is a supplement that varies from country to country, depending on culture and institutions.*" (Ibidem, p. 8-9) This opportunism is the undertone in much of TCE. According to TCE, actors are real bad guys: "*Opportunism refers to self-interest seeking with guile. The image of human agents is grim: they are assumed to lie, steal, and cheat, as well as to use more subtle forms of deceit.*" (Ibidem, p. 8) Yet people do not have to be opportunistic all the time. It can be the case that some individuals have sufficient personal integrity: "*where personal integrity is believed to be operative, individuals [...] may refuse to be part of opportunistic efforts.*" (Ibidem, p.16) Due to the fact that people are aware of opportunism of other people, they will choose to form the transaction in such a way that the space for the other to employ opportunistic behavior is reduced.

The always hidden presence of opportunism is a core theme within TCE. When two firms enter into a transaction with one another, both parties are trying to optimize their profit with respect to their costs. Because it is most likely that the other only is concerned with his own interest and will possibly employ opportunistic behavior, people will try to minimize this risk and also the costs that are concerned. If an actor had unbounded rational faculties, he could take all possible outcomes into account during the period of contracting. This results in a perfect contract. Unfortunately, people are limited in their ability to predict possible outcomes of transactions. This means that governance structures have to be devised in order to reduce the risk of opportunistic behavior: "*The central question of economic organizations is to "devise contract and governance structures that have the purpose and effect of economizing on bounded rationality while simultaneously safeguarding transactions against the hazards of opportunism.*" (Ibidem, p.6) The safeguard that governance structures offer, is that the loss the breaking party will endure will surpass the profit of breaking the contract, without the need for a court to settle the differences or impose corrections later on.

Take for example a situation in which a manufacturer produces car doors for a large car factory. This manufacturer has to invest product specific machines and buildings to construct the doors. The manufacturer therefore requires minimal sales to make a profit. A premature breaking of the contract will render the specific investments worthless which the manufacturer will try to avoid. On the other hand the car factory needs car doors and can only buy these doors from specialized firms of whom only one may exist. When the car door manufacturer breaks the contract, the line of production may stop for weeks resulting in enormous costs. Therefore, for both parties it is costly to break the contract. This form of governance in which both firms will loose from contract breaking is called bilateral governance. It can also be the case that only one of the parties will loose from contract breaking. This is called unilateral governance: "*Two kinds of differences between (...) [unified governance and bilateral] governance (...) can be distinguished: differences in incentives and differences in the monitoring, controlling and auditing apparatus.*" (Ibidem, p. 9) The different kinds of governance result in different costs: "*These various governance structures are also distinguished by differential costs: real costs are associated with negotiating, writing, monitoring and enforcing contracts, and, in addition, governance structures that shut out the market suffer from a loss of incentives*." (Ibidem, p. 7) The TCE actor tries to optimize the costs with respect to the expected profit. Many of the costs are concerned with governance structures. About transaction costs itself TCE says the following: "*Williamson follows Arrow's definition of transaction costs as the costs of running the economic system, and as such they are equivalent to 'friction' in the physical sciences. In more detail, ex ante transactions costs are the costs of drafting, negotiating, and safeguarding an agreement. Ex post transaction costs include: (1) the maladaption costs incurred when transactions drift out of alignment with requirements, (2) the haggling costs incurred if bilateral efforts are made to correct ex post misalignments, (3) the set up and running costs associated with the governance structures (often not the courts) to which the disputes are referred and (4) the bonding costs of effecting secure commitments*." (Dietrich, 1994, p.21) Although different kinds of costs can be minimized, this happens in a depersonalized domain. Despite the fact that opportunism is somewhere ruling the game behind the curtains, the actor and its characteristics are completely invisible within this explanation of TCE

2.2.2 Opportunism and loyalty

Maximization of profit is the core of many economical theories. According to TCE this maximizing is not an easy game. Other actors are opportunistic as well and all actors have faculties that are boundedly rational. With regard to his limitations an actor tries to behave as smart as possible by using governance structures that reduce opportunism. "*The central question of economic organization is to "devise contract and governance structures that have the purpose and effect of economizing on bounded rationality while simultaneously safeguarding transactions against the hazards of opportunism.*" (Noorderhaven, 1993, p. 6) According to TCE, actors aim at self-fulfilment. In TCE terms this means that actors only try to fulfil their own goals. These goals are related to making profit in the present or in the future. This may be achieved at some other one's cost by deceit or racket: "*Opportunism refers to self-interest seeking with guile. The image of human agents is grim: they are assumed to lie, steal, and cheat, as well as to use more subtle forms of deceit*." (Ibidem, p. 8) Some societies may have a basic form of trust as a result of culture or institutions (e.g., the church), but (inherent) human opportunism forms the invariant core: *"[O]pportunism forms the invariant core of economic actors; trustworthiness is a supplement that varies from country to country, depending on culture and institutions*." (Ibidem, p 8-9) Actors are aware of the opportunism of others. They actively try to arm themselves against the hazards of other people's opportunism. This opportunism does not have to lead to a breaking of the contract. Even if you know that your partner is opportunistic, you do not know whether he/she will take advantage of this opportunism: "[S]*ome individuals are opportunistic some of the time and differential trustworthiness is rarely transparent ex ante"* (Williamson 1985, p. 64). Because it is hard to predict who will behave opportunistically at what time, opportunism is always possible. This means that adequate measures to refrain from its negative effects are always needed.

The impression one gets concerning opportunism within TCE can be made more clear by using Nooteboom's distinction between *the space for* and *the inclination to* opportunism. TCE does not take inclination to opportunism into account, because "*[i]nasmuch as a great deal of information about trustworthiness or its absence that is generated during the course of bilateral trading is essentially private information (...) knowledge about behavioural uncertainties is very uneven. The organization of economic activity is even more complicated as a result"(Williamson 1985, p. 59).*" (Klos, 2000, motivation in his abstract) This implies that TCE is a simplified theory. Knowledge about behavioral uncertainties is asymmetrically distributed. As a result the organization of economic activity is even more complicated. The simplification TCE makes, is that it assumes that the inclination to opportunism is always large. Because keeping track of the trustworthiness of all people from past, present and future is complicated, to say the least, TCE assumes the same trustworthiness for everyone. In this way equal trustworthiness is the same as a maximum inclination with regard to opportunism. Again we see that the space for opportunism is the critical factor. Building safeguarding mechanisms is useful, because it minimizes the space for opportunistic behavior. TCE takes loyalty as the counterpart of opportunism. However, loyalty is negatively defined with regard to opportunism. Someone who does not behave opportunistically, is considered to be loyal. This is especially true with regard to behavior within the organization. TCE assumes that within an organization all employees are considered to be completely loyal.

2.2.3 Bounded rationality

TCE reserves the notion of bounded rationality for both the lack of worldly knowledge of the future and for the presence of limited computational facilities of the processing mechanism, i.e., the human cognitive system. Bounded rationality makes it impossible to have a good general view on all possible situations that can take place. As a consequence it is impossible to devise a contract that incorporates safety in such a way that all possible situations in which one can become victim of opportunistic behavior are foreseen. Therefore, specific investments are risky. "*Because actors are assumed to be opportunistic, ... investments [in highly specific assets] will, in the absence of adequate safeguards, lead to costly haggling and expose the party incurring them to the risk of expropriation. If actors had unbounded rational faculties, a complete and contingent contract safeguarding specific investments could be written. However, TCE assumes bounded rationality.*" (Noorderhaven, 1993, p. 6) Actors, with their limited processing capacities, try to achieve as much as possible. No unwillingness is said to be in play: "*Bounded rationality* *refers to the fact that people (agents) are intendedly rational, but only limitedly so.*" (Klos, 2000, p. 11) This bounded rationality results in uncertainty. Sometimes the actors just hope that the result of a transaction will be positive. This assumes some kind of trust, which we will discuss later.

Because the set of all possible future states cannot be predicted, it is the goal of governance structures to prevent some events from taking place. By deselecting bad options this makes the set of future states smaller. The assumption of opportunism will leave the non-relevant parts of the set of possible future states out of consideration. This has a positive effect on the bounded rational capacities, that is to say that more can be achieved "*The central question of economic organization is to "devise contract and governance structures that have the purpose and effect of economizing on bounded rationality while simultaneously safeguarding transactions against the hazards of opportunism.*" (Noorderhaven, 1993, p. 6) A small point of criticism can be made, here. The concept of bounded rationality is hardly elaborated upon. There are two interpretations of the concept of bounded rationality: an ontological and an epistemological one. The ontological interpretation refers to the impossibility to determine from a point in time onwards the course of events or the development (i.e. to foretell the future). The amount of rain that will fall in two months is not predictable, due to the dynamics, complexity and layering of the material world. Epistemological rationality refers to the boundedly cognitive components, both in the architecture and in the representations of cognitive actors. So, even if it was possible to know the course of events in advance, it would be impossible for us humans to comprehend everything due to our limited rational faculties. Most discussions on bounded rationality in the economical sciences - and also in TCE - are concerned with the ontological interpretation, whereas Simon (1960) who introduced the concept of bounded rationality emphasized the epistemological, i.e. the cognitive, interpretation. Our constraints as humans can thus be found, both in the gathering and in the processing of information. Even if we can learn, our cognitive system keeps playing us tricks, because of its cognitive architectural structure.

2.2.4 Learning

As in most theories, TCE assumes people to be capable of learning. Experiences with partners will affect the trustworthiness of partners and will have learning effects on the future behavior of people. Furthermore, learning can enable humans to use more elaborated forms of analysis. Within TCE learning comes in just one variety, i.e. learning in the sense of acceleration or optimization: "*individual and personal trust relations evolve at the interface between supplier and buyer*" (Ibidem, p. 16) This is sometimes expressed as a form of saving: "*Additional transaction-specific savings can accrue at the interface between supplier and buyer as contracts are successfully adapted to unfolding events and as periodic contract renewal agreements are reached. Familiarity here permits communication economies to be realized: specialized language develops as experience accumulates and nuances are signaled and received in a sensitive way. Both institutional and personal trust relations evolve*." (Dietrich, 1994, p. 28) The prediction of the behavior of a partner as a result of his different attributes is said to become more accurate and the resulting search space will be less complex admitting a better result. It is but one of the two interpretations of bounded rationality that profits: more information is gained about a person, so a better picture can be attained. The idea of a better analysis based on the same data is never suggested. This aspect of learning, namely incremental or creative learning, is not addressed by TCE.

2.2.5 Trust

In TCE trust and opportunism are closely related. Williamson says: *"[O]pportunism forms the* *invariant core of economic actors; trustworthiness is a supplement that varies from country to country, depending on culture and institutions."* This means that the basic perspective on actors is one of grasping the opportunities in a seemingly unscrupulous way. If two options are presented or encountered that require a choice, the one with the highest expected profit and the least risk is necessarily chosen. However, actors are aware of their own and others opportunism. All parties seek for safeguards against the threat of opportunism. This opportunism does not mean that in all cases you or your partner break the relation. Even if only a few safety measures are taken, a transaction can be fruitful for both partners in case they keep the contract instead of a betrayal of the other. So, even if you know that your partner is opportunistic, it does not mean that he will be opportunistic in reality. *"[S]ome individuals are opportunistic some of the time and differential trustworthiness is rarely transparent ex ante"*(Williamson 1985, p. 64)

As indicated TCE mainly talks about opportunism and rarely about trust. However, opportunism and trust are two sides of the same coin; if you know the one, you also have the other. Because, according to TCE, opportunism is more fundamental than trust, it seems that TCE strictly follows the Hobbesian proverb “homo homine lupus est”. However, even in this negative perspective on human nature, it is assumed that opportunism as a label for certain kinds of behavior basically is a decision making situation. If I choose for the new more profitable partner, even though I have a ten-year lasting relation with my present partner, a decision making process has taken place and a decision outcome is cognitively calculated and realized. In our view TCE’s perspective on trust and human behavior can only be true if the following three assumptions are taken into account. First, all information necessary for a choice is present and can be evaluated in a (quantitative) model. If information is not present or cannot be quantified it is not essential and therefore not relevant. Second, the information processing limitations of an actor are not relevant in the case of making a choice, whether the information is complete or not. Third, the (cognitive) mechanisms of actors to deal with opportunism and trust are adequately known and the part(s) that are not known are irrelevant. We will discuss the three assumptions in some detail and will come up with a supplement to the missing part related to individual cognitive processes later.

The presence of complete information is an assumption that can be understood from the question or issue of how to define **in**complete information and how to deal with it. It is well-known that within TCE incomplete information, non-transparency of the problem domain and unpredictability as a result of the limited overview of a person in relation to the problem domain can be partly called “bounded rationality”. However, complete information does not mean knowing everything on all aspects, but complete has to be seen in relation to the decision to be made. If an actor is considering to have a relation with another party - meaning to buy or to supply - then that actor makes judgements and uses information and knowledge, that in restricted time lead to a shift toward another partner or a continuation of the present relation. This means that saying that information should be complete is making a straw man within TCE that can easily be set to fire. Complete information does not mean information of everything, but it means information in relation to the situation. What do I care that on the other side of the globe someone considers to start a new service, unless of course this information affects my own decision situation. However, because the situation is always in relation to the actor - the decision maker - it means that estimations, personal histories and preferences come into play. This process causes much uncertainty and here trust comes in. Numan (1998, p. 30) says that trust is a way of reducing a feeling of complexity, uncertainty and risk. Numan (p. 31) continues to say that trust is anticipating the future by assuming that the future is certain. In this way people try to reduce uncertainty about the future. Trust is an expectation with regard to certain aspects of the subject and other’s behavior. Within TCE opportunism may be the basic attitude of actors, it also implies that actors have to take some things for granted. For example, if I substitute my present partner for a better one - promising me more profit, a more durable relation and little risk - which in terms of TCE may be called an act of opportunism, I still have to trust that my new partner is alive tomorrow. Therefore, opportunism cannot possibly mean to act completely under the influence of "the here and the now". Many things always have to be taken for granted, which means that I have to trust on many things even in an act of so called opportunism.

Another point of interest is that trust suffers of what can be called the information paradox. Trust is based on information, but trust is only relevant if information is not complete, that is to say if there is uncertainty. Trust means that one exposes oneself to risk. One is not only uncertain regarding the temptations or the pressure with which one's partner might be confronted, it is also not ethically correct to limit a partner's freedom to act and to treat him or her as deterministic and static. Partners are as free as you consider yourself to be, and the trust you have in others might reduce the feeling of uncertainty, it will never eliminate uncertainty.

As a second assumption TCE supposes or - perhaps better - should suppose that opportunistic or trustworthy decisions are the result of information processing activities of actors. In cognitive terms actors as information processing (or cognitive) systems have a cognitive architecture, have representations and execute operations on these representations leading to actions, for example continuing a relation or starting with a new partner. In this cognitive process the architecture, for example human memory, has limitations. Well known is Miller's (1956) finding that our short term memory system has a limited capacity of seven plus or minus two meaningful elements. Also the representations we work with are limited in the sense that a language-like representation cannot easily be integrated with an image-like representation. In the same way the operations are limited in use and scope. Though we want to forget certain information, it is nearly impossible to make that cognitively happen. The operation of forgetting - or deleting - is not freely available and applicable, even if we strongly want certain things to be forgotten. We will return to the issue of human information processing in section 5, where we will suggest a cognitive supplement to the rather empty actor model of TCE.

The third assumption regarding opportunism or trust in TCE is related to the cognitive mechanism that realizes what we call trust. What we know is that trust does not suddenly come into being. It needs a history with a reliable background. Depending on the specific history, trust is built up and broken down. Numan (1998) indicates that trust can be based on external sources like the opinion of another person. It can also be based on empirical facts and it can be based on representations of the actor. When trust is given, it is not relinquished lightly, although drastic events indeed reduce trust. As Numan says, people's trust is something we can regard as quite stable until something happens the person had not expected, after which this trust in certain things has to be re-evaluated. New relations of trust come into being when an actor interacts with a person in a very complicated way, so the actor has to make a choice between different options that involve various kinds of risk. An actor can choose between three different ways of dealing with such a problem. The actor can look for help, he can act by trial-and-error and he can use so-called inference mechanisms. In the latter case trust comes in. Inference mechanisms or reasoning procedures are necessary for every estimation or anticipation of what a relevant and adequate future action could and should be. The problem is that the cognitive mechanisms that enable trust are invisible, at least in TCE. They have to be stipulated. Numan (1998) defines trust in cognitive way as follows. "*Trust is a mental action. This action is an expectation which person A has of an actor B – that actor B will act positively toward the goals which the trusting person A has. In this, the actor B, who has to be trusted, has the freedom to harm the trusting person A. The expectation is based on incomplete evidence.*" (Numan, 1998, p. 39)

If we leave out the reciprocal situation of actor B toward actor A, the definition basically is about cognitive activities. Formulated in cognitive or information processing terms the mental action, namely expectation, is a computation on mental representations, where representations are the expression of the various contents of the knowledge actor A has.

Because actor A has knowledge before the decision of continuation or change of the relationship takes place, there also has to be some kind of base trust. It is highly unlikely that for every possible partner a kind of mental list is stored and updated in the cognitive system of actor A. This assumption of a base trust does not imply that every actor has the same level of base trust. Individual variation is the norm and not the exception. Making a distinction between high and low levels of base trust is also possible. A high level of base trust can be called loyal and a low level can be called opportunistic. High levels are resistant toward temptations from other attractive partners, but low levels are not. Base trust will of course be affected by experiences. Someone who has frequently been cheated will have low base trust and will, even in the case of a new relation, act cautious toward the breaking actor. This may result in future opportunistic behavior by the cheated actor. In a cognitive system the memory parts are responsible for the storage and updating of past and present knowledge and estimations concerning the future. Therefore, general trust implies a dynamic base trust (DBT) and momentary trust (MT). DBT is not a constant, but a slowly changing variable, that has to be established in time. MT relates to the present trust and is updated with every new experience with regard to the partner with which a relation right now exists.

With regard to trust in general and because trust has so many synonyms, Numan (1998) proposes the amount of evidence in relation to a decision making situation as a demarcation criterion. Evidence is the estimation and expectation regarding things that have happened, happen and will happen and can come from internal memory and the interpretation of external information. If actor A has inconclusive evidence we call his state of mind trust. If actor A has no evidence at all and (still) vehemently believes that certain things are relevant or will happen, we call his state of mind faith. When actor A thinks he has conclusive evidence there is confidence. Therefore, whether an action is based on faith, trust or confidence depends upon how a person experiences and thus represents the evidence.

At this point some important individual key concepts of the TCE actor that can be found in the literature have been discussed. The sum of these descriptions should result in a solid and complete enough model of a TCE actor. One may reasonably expect that this model of an actor largely corresponds to one's own idea of an economically behaving human being. Furthermore, one may expect TCE to be based on descriptions that fit with the cognitive literature. To be blunt, we doubt whether this is true. Do we now have an adequate model of an actor within TCE? And is this actor cognitively plausible? Or, to go one level deeper, what kind of actor does TCE take for granted and what are the components of this actor? In the next section we will try to answer these questions.

3. Plausibility, levels of description and kinds of actors

3.1 The cognitive psychological plausibility of a TCE actor

From the discussions regarding TCE and the role of the actor it can be concluded that first actors are completely loyal towards individuals within the organization, but also completely opportunistic with regard to other actors outside the organization. Second, within TCE, learning means acceleration and not innovation or creativity. How this process of acceleration takes place is not explicitly mentioned. Third, organizations and companies (firms) try to optimize their behaviour, but optimization itself is realized externally by the market and not by the companies (actors) themselves. Fourth, the assumption within TCE is that opportunism is not the result of calculated individual weighting and individual knowledge about uncertainties. Knowledge and behavior are not uniformly present in a company. TCE, therefore, assumes that the propensity towards opportunism is always large. It also is too difficult to keep track of (individual) trust. Fifth, although bounded rationality is taken for granted, it is not much elaborated. Also a confusion exists concerning the precise interpretation of bounded rationality: an ontological and a epistemological/cognitive interpretation. The ontological variant says that gathering information is limited because it is impossible to know the complete state of the world over 10 days or 5 years. The complexity is beyond imagination. The cognitive interpretation says that, besides the fact that it is impossible to predict the complete future state of the world, the human information processing system has its limitations in cognitive architecture, representations and operations on representations.

Although TCE may be consistent in terms of its own economic assumptions, it has to be questioned whether the fundamental notions with regard to plausible human actors are adequate. TCE implicitly accepts the limitations of its approach with regard to human actors, but no consequences are visible with regard to this sensitivity. This is not a negative point of TCE as such, because the requirements of an economic theory of behavior are less complex than the requirements for any general theory of human behavior. However, the assumptions concerning an actor should not be in contradistinction to what we normally expect from a cognitively plausible human actor. In the foregoing we already interpreted TCE much more favorable than can directly be inferred from the actual statements of TCE concerning actors. TCE is very vague about what actors really are and how they function and process. The position we want to put forward, here, is that without a plausible (cognitive) actor theory, TCE might be incomplete and therefore its predictions concerning economic behavior might be unrealistic. Individual characteristics such as opportunism, bounded rationality (in its cognitive interpretation), learning and trust are not founded on a cognitive theory. This is odd, because if TCE realizes models in which no constraints are included for components of cognitive actors, TCE might become largely implausible itself. It is as if characteristics of molecules are not bounded by some physical theory of particles in which electrons and bindings are accounted for. If discussions about trust, learning and opportunism in behavior are not rooted in so called “minimal” cognitive models, then Newell & Rosenbloom’s (1981) general remark can also be applied to TCE. "*By delineating boundaries on the set of all legitimate models, constraints provide a means by which knowledge can be transferred to later more correct models of the task. Constraints can be formulated in terms of either the behavior of the model or the structure of the model. Behavioral constraints are the weakest and are provided in profusion by the results of psychological experiments. Structural constraints are much stronger; they limit the space of structures that may be inside the box. At their most powerful, structural constraints are capable of ruling out large classes of architectures.*" And, we would like to add, large classes of economic behavior.

It should be clear from the above that the big omission within TCE, but also the big challenge for TCE is a description of underlying processes and structures of the actor. Although properties and characteristics of actors are mentioned, they are not rooted in a cognitive model or architecture of the actor. It should be kept in mind that TCE is an economic theory. Such a theory tries to provide insight of the economic world by only taking into account the relevant parts of economic behavior. The more seemingly irrelevant and disturbing parts are included into the theory, the more complex the theory becomes. It seems that TCE tries to accomplish a simplified view by assuming that every normal individual has adequate knowledge about his or her self. Everybody or every well educated human being knows what it means to have trust, to be loyal or to be rational, so why bother about more subtle components. They only make the models more complex. This is a useful starting point for any theory of behavior, provided that the basic notions about human actors and human behavior are self-evident. We believe that this position is not as powerful as is suggested. The problem with taking what everybody, in a standard way, knows about himself as an actor is that there are many degrees of freedom and many variances. To which degree are the monk who is sacrificing his life to become a martyr or the soldier giving his life for his country representative for the normal behavior of actors. Everyone would argue that these examples are extreme, but isn't the reverse argument applicable for the so called uniformity or non-differentiation of actors within TCE. An actor model can not be constituted out of a so called uniform actor, complemented with what everybody knows of himself. We would like to argue in line with classical cognitive science, formulated by Newell & Rosenbloom, that it is not sufficient for any theory of human behavior to state that "everybody knows what it is to be human being".

Given the before mentioned analysis two lines of reasoning can be followed. First it could be argued that TCE is debatable or could be ignored because of its negligence of human information processing components. We would not argue in favor of this line of reasoning. Instead we would like to advocate another line of reasoning in which attempts are made to substantiate economical models of human behavior by providing the models with plausible cognitive components of the actors. What does it mean to model and implement an acceptable actor model and what cognitive components can be proposed? This is in line with the work of Simon, March and others on the behavioral theory of the firm. We will try to do this in the remainder of this article. However, before we can elaborate the cognitive actor components we first have to discuss two issues. The first regards the various levels of description of an actor, the second regards the kinds of actors and its components that we can discern.

3.2 Levels in describing kinds of actors

The idea of levels of description or explanation for a (cognitive) system has been worked out most elegantly by Dennett (1978, 1987, 1991) who distinguishes three independent levels, the physical stance, the design (or functional) stance, and the intentional stance. Other authors (Newell, 1982, 1990; Pylyshyn, 1984) have given similar accounts in which, however, the number of levels vary. In the following paragraphs we use Dennett's explanation of levels (1978, 1987).

The physical stance explains behavior in terms of physical properties of the states and the behavior of the system under concern. For its proper functioning the human organism requires a complex interaction between its parts and with the external world. The central nervous system, and the endocrine system are there to transmit information that reveals the state of one part of the system to other parts. We can also mention the transmission of currents in the synaptive system of neurons. In cognition, this stance is the endpoint of successful ontological reduction.

The second level of explanation takes the point of view of the functional design of a system. The behavior of a system is conceived of as the result of the interaction between a number of functional components or processes. The physical structure (architecture) of the system is not explicitly taken into account, although it may impose constraints on the behavior of the system. The capacity limitations of human memory will, for instance, impose a boundary on the complexity of decision making.

In the third place Dennett distinguishes the intentional stance. Complex behavior that is adapted to the prevailing circumstances, according to some criterion of optimality is said to be rational or intelligent. A behaving system to which we can successfully attribute rationality or intelligence qualifies as an intentional system. It is not necessary for a behaving system to 'really' possess rationality or intelligence, as long as the assumption allows us to correctly predict the behavior of the system on the basis of our knowledge of the circumstances in which the system is operating.

One may deal with this three-cornered distinction in two essentially different ways. Dennett (1978) has taken this distinction in a strictly instrumentalistic way, claiming only a pragmatic validity. Summarizing his position twenty years later, he wrote: "*As I have put it, physical stance predictions trump design [or functional] stance predictions which trump intentional stance predictions - but one pays for the power with a loss of portability and a (usually unbearable) computational cost*." (Dennett, 1998, o.c., p. 119) In contrast to Dennett, authors such as Fodor (1975), Newell (1990), and Pylyshyn (1984) assign an ontological significance to each level. According to these authors the higher levels introduce emergent qualities into human behavior that make no sense if we maintain an instrumentalistic point of view.

3.3 Components of actors and kinds of actors

The distinction in levels of description is meant to analyze complex systems such as actors. If we look at TCE we see that the dominant level of description is the intentional level. Trust, loyalty and opportunism within TCE are all at an intentional level. It should be clear that even with a cognitively plausible actor, behavior of economically normal acting human beings at an intentional level must be given. This means that an actor model should not include components that are in contradistinction with the real life perspective we have of people. That is good for a start, but it is not enough; it is a necessary, but not a sufficient condition. An actor also has to be plausible at a functional level of description; that is an actor model has to be congruent with what is known from received cognitive science. This means that minimal requirements have to be formulated and implemented for actor models. The fundamental structure of an actor should be consistent and also minimally complete. Theoretical predictions at the intentional level of economic behavior should not be contradicted by the functional architecture of the actor. For example, if it is stated in TCE that an actor stores information or learns, at the functional level the actor should have a memory and a learning mechanism. A memory module, therefore, is a necessary component of any (intelligent) actor.

Within the cognitive literature a extensive list of cognitive components has been formulated (Posner, 1989). It is not our intention to discuss the complete list, but some aspects are relevant in any actor model. Components that can be found include memory (ies), information processors with limited capacity, memory content consisting of various representations, such as propositions, icons, images, episodes, scripts or procedures. Furthermore, there are sensors and effectors (motor components), interacting components, leading to new functions, such as a communication channel and (personal) goals, which are present in the form of representations, for example the maximization of perceived profit within the TCE model.

The above mentioned components can be combined in various ways. If we consider an organization to be a multi-actor system, various compositions may be relevant in the description of various kinds of actors in the organization. This is a basic assumption in our position, that in the end any organization consists of a multitude of actors. An organization functions because of the performance of the various actors complemented by communication channels and coordination mechanisms (H. Gazendam, 1993).

Integrating the elements above, actors can be discerned regarding the presence or absence of the following components: a) *perception*, b) *interaction* (including learning in the sense of habit formation), c) *representation and interpretation* (including learning in the sense of chunking and creation) and d) *autonomy and self-consciousness* (H. Gazendam & Jorna, 1998).

With *perception* we mean that a system must be able to accept input in a general sense. This input may include visible, audible and tangible stimuli and the accepting system may vary from a lobster to a human being or even a computer system.

Interaction is the process by which a system has contact with its environment. Stimuli as input in the system lead to output in the sense of responses. The reaction patterns of the system may result in learned behavior, that is to say that habits are formed.

A system that internally symbolizes the environment is said to have and use *representations*. Representations consist of sets of symbol structures on which operations are defined (Jorna, 1990). Examples of representations are words, pictures, semantic nets, propositions or temporal strings (Kosslyn, 1980; Anderson, 1990). A representational system learns by means of chunking mechanisms and by the creation and manipulation of symbol transformations.

A system is said to be *autonomous, self-organized or self-conscious* if it is able to have a representation of its own (physical and conceptual) position in the environment. This means that the system has self-representation. An autonomous system has reconstructing representational interaction patterns.

The four aspects together result in an actor hierarchy. An actor that only has perception is at the lowest level and can not be called an intelligent actor, whereas an actor with self-organization, including perception, interaction and representations, is the highest level. This last form is what we regularly call an actor that is reflective, intelligent and thoughtful. Human beings are good instantiations of intelligent actors. Computers are said to have representations, but not self-organization.

Not every actor is intelligent, but every intelligent system is an actor. The above described classification in perception, interaction, representation and autonomy can be related to a qualification of actors. First we subdivide actors in response function systems, representational systems and representational response function systems. The use of the term system implies that we consider an actor to be a coherent whole, consisting of several components, for example motor parts, sensory parts, including perception, and cognitive parts. The parts will not be discussed in detail, here (see Posner, 1989; Newell, 1990). Second we make a distinction in single actors and multiple actors Concerning multiple actors the surplus component is a (complicated) communication and coordination mechanism, that realizes the interaction of actors. This mechanism has to be intelligent and should, in an "ultimate analysis", also be incorporated in a human system.

We start with a cohesive, structured and organized entity. This entity operates in an environment, but no specifications of its operations are given. In a sense this entity is an actor, because it is self-contained, strives toward continuation and, looking at the actor characteristics, it has perception and interaction including the possibility of learning in the sense of habit formation. We emphasize that this actor does not have internal representations. Its cognitive domain is absent or empty. We call this actor a response function system (RF-system), or Actor I (see figure 1), and it can be compared with the ant in the sand (Simon, 1969). In discussing complex behavior of systems Simon stated that the behavior of an ant on the sand can be called complex, although not intelligent, because the behavior is a function of the complexity of the sandy irregular environment that the ant has to cross.

In the second place we can conceive of another actor that we call a representational system (R-system). This actor has representations and is able to project external events internally into its cognitive domain. We call this Actor II (see figure 1). This representational system has representation, autonomy and perhaps perception. The interaction is problematic, that is to say that there is no device that semantically interprets causal input and output. As far as we talk about interaction it is rather low level reaction to stimuli.

The third possible interpretation of an actor is the representational response function actor (RRF-system). This actor incorporates a really intelligent, interactive and cognitive system. We call this Actor III (see figure 1). This actor is able to perceive, to interact, to represent and to be autonomous. Cognitive processes include symbols, operations and semantic interpretable response functions. RRF-systems behave on the knowledge level, as Newell called it. "*There exists a distinct computer systems level, lying immediately above the symbol level, which is characterized by knowledge as the medium and the principle of rationality as the law of behavior.*" (1982, p. 99) Newell is proposing this knowledge level for natural (humans) as well as artificial (computers) intelligent systems.

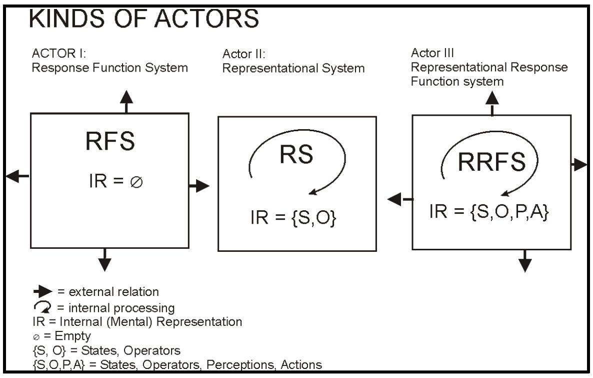

Figure 1: Overview of Kinds of Actors

The actor equipped with the integration of representations and responses has knowledge. "*Knowledge*", says Newell, "*is whatever can be ascribed to an actor, such that its behavior can be computed according to the principle of rationality.*" (Newell, 1982, p. 105) The principle of rationality is expressed in the belief that an acting person will undertake those actions by which his goals are reached. These actors are submitted to what Simon (1947/76) called: "bounded rationality".

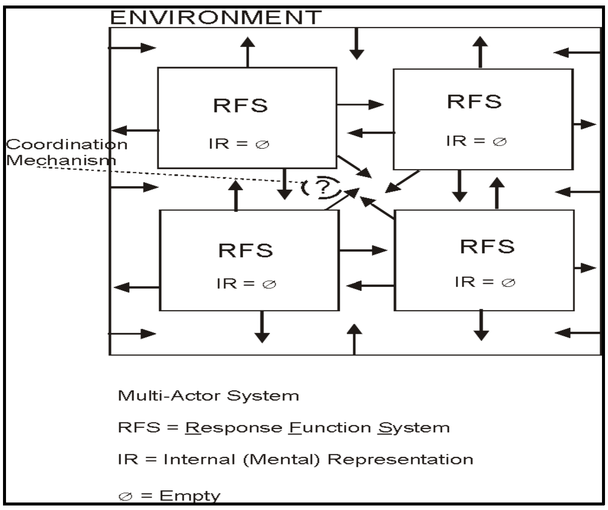

Figure 2: RFS (Multi-actor System)

The hierarchy of single actors returns in the composition of multi-actor systems. In the first place it is possible to have a multi-actor system in which the actors are response function systems. The situation is comparable to the single actor system in that the actors do not have internal representations. In this circumscription of multi-actors other actors are considered to be parts of the environment. The only difference is that there is a proximal and a distal environment. The system borders define an inner and an outer area. The actors all have perception and interaction. To take up the example of Simon's ant we are talking here about a group of ants perceiving and interacting with each other. Coordination is absent or only defined in terms of reactions to behavior of other actors. To make a provocative statement, we state that although organizational theory speaks about the coordination of multiple intelligent actors, in practice the actors are mostly defined as response function systems (see figure 2), that is to say as Actor I (Klos 2000)

In the second place we have a multi-actor system consisting of representational systems (Actor II). Every actor has internal representations in the sense of symbol structures and operations. Interaction is only possible if the symbol structures are similar, that is to say in the form of strong codes, such as notations (Goodman, 1968). It is of course doubtful whether interaction between the actors is semantically meaningful. In discussions about social cognition the issue of semantic interaction is ticked off, but not resolved. Intelligent coordination without communication in terms of notations is hardly handled (see figure 3).

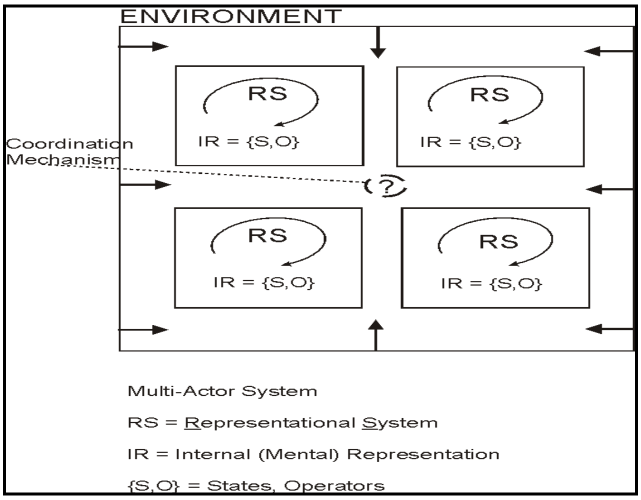

Figure 3: RS (Multi-actor system)

In the third place there are the representational response function systems in a multi-actor situation. The actors perceive each other and react to each other in a semantically rich and intelligent way. Each actor has perception, interaction, representation and autonomy and manages to integrate this into the organization as a multi-actor system. A collection of human information processing systems - the organization in a multi-actor perspective - is an example of multiple representational response function systems. This is the situation of so called organizations in practice. They consist of actors in the sense of actor III (see figure 4).

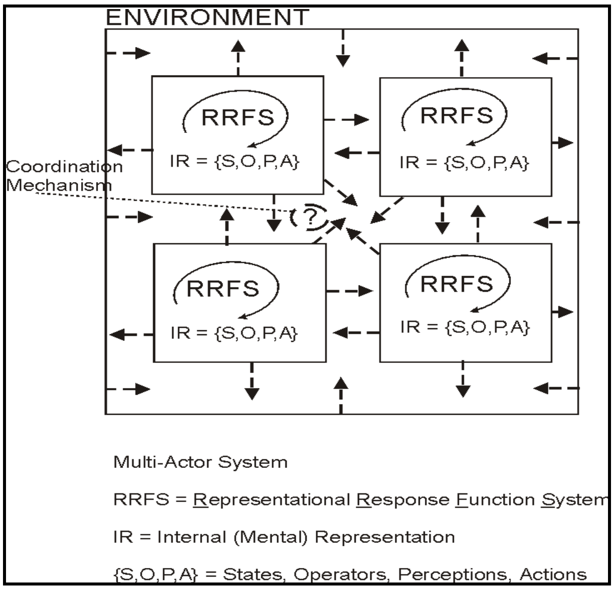

Figure 4: RRFS (Multi-actor system)

In the fourth place a combination of several kinds of actors is possible. Various combinations of actors can be distinguished such as multi-actor systems consisting of actors I and III, actors II and III, and actors I, II and III.. A combination of actors I and II seems difficult because at last one of the actors in a multi-actor system should consist of autonomy and self-organization. Other multi-actor systems, that partly coincide with actors I, II and III, will consist of natural and artificial intelligent actors. Under the influence of developments in ICT more and more artificial actors will behave as actors III, which to many people now seems to be horrifying (see figure 5).

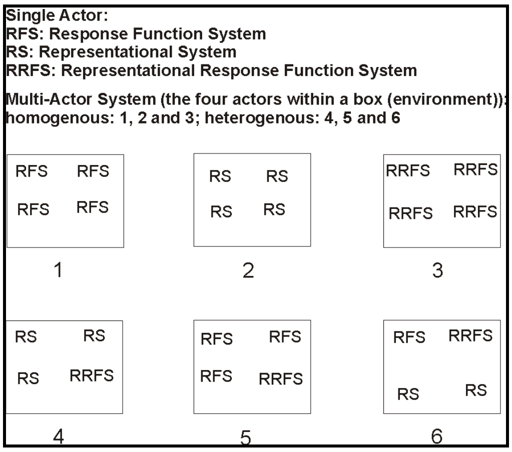

Figure 5: Heterogeneous and homogeneous Multi-Actor Systems

The only meaningful incorporation of knowledge, cognition and representation in organizational theory is in a representational response function system. This holds for a single actor as well as for multi-actor systems. As already indicated, this does not mean that all actors in such a multi-actor system have to be RRF-systems. Some actors in the multi-actor system may be

RF-systems or R-systems, but at least one of the actors has to be a RRF-system. The interpretation of an actor or a multi-actor system as a RRF makes it possible to introduce cognition, knowledge, interpretation and symbol/sign manipulation concepts. [[This combination of terms refers to cognitive science (Posner, 1989), to knowledge management (Jorna, 1998) and to semiotics (Michon, Jackson and Jorna, 2002). Cognition and knowledge will be discussed in section 4. We will end this section with some remarks about coordination and communication and the entities that perform this job: the signs and symbols.

The coordination and communication between actors is expressed in signs and symbols. This implies that actors within organizations also are semiotic entities. With semiotic we mean that an organization as an artefact, a construct or a representation, basically is a sign type or sign token. The interesting point of looking at actors and organizations as semiotic entities can be found in the different sorts of signs that turn up in the communication and information structures of the various actors. In semiotics it is normal to distinguish signals from signs and to subdivide signs in icons, indexes and symbols. Icons emphasize the similarity aspect, indexes the contiguity aspect and symbols the conventional aspect of signs. In relation to a deeper discussion of characteristics of actors of type III - the RRF systems - semiotics gives a conceptual apparatus to deal with knowledge, communication, representations, symbol structures, interpretation and meaning (Nöth, 2000; Michon, Jackson, Jorna, 2002). This can be applied with regard to external communication and with regard to the internal mechanisms of actors. It is impossible to discuss external communication without the conceptual framework of semiotics. However, that is not the perspective in this article. Here, we limit ourselves to the relevance of symbols in the classical cognitive science approach. We will return to these cognitive actors, their limited implementation in artificial intelligent architectures and TCE as part of a more general discussion of kinds of actors.

4. Intelligent actors and architectures for cognition

As stated in section 3.1 an actor not only has to be plausible at an intentional level, the actor has also to be plausible at a functional level of description. That is, an actor model has to be congruent with what is known from evidence within cognitive science. This means that minimal requirements have to be formulated and implemented for actor models. The fundamental structure of an actor should be consistent and minimally complete. Just like the parts of a car engine without a sparking plug fail to generate their intended function at a higher level (i.e. a running engine), all the components of our actor at the functional level are essential for the behavior at the intentional level. Therefore, the completeness at the functional level is determined by behavior we desire at the intentional level.

In looking for a fundamental structure for a cognitive plausible TCE actor (a TCE-COG actor) we can find many models in cognitive science. Since the completeness of the functional level is determined by the intentional level, it is at forehand hard to see what model should be used. In the categorization of (aspects of) actors we have seen that the most complex actor is the representational response function actor. This actor includes all the components that can be found separately by the other, less elaborated, actors. The cognitive model we select for the TCE actor should contain the structure of or allow for a RRF system. The idea behind this is that TCE tries to model economic behavior of complete humans and not of so called stripped actors. TCE does not model the economic behavior of ants (RF-systems) or of groups of solipsists (RS-systems). A RRF system also forms a bridge between the functional and the intentional level, which should facilitate the choice of the model. The cognitive theories that claim to address the complete spectrum of human cognition are the Architectures For Cognition (AFC). The development of these architectures for cognition is inspired by the development of the computer. A small detour in the direction of the developments of AFC will make clear what kind of cognitive model and architecture might do the job for the presently incomplete TCE-actor.

In the past various attempts have been made to simulate human cognition on computers. In the seventies hardware and software became so sophisticated that realistic cognitive actor models could be designed and implemented. Conceptually this was realized by Turing's theories about computational functions (Millican & Clark, 1996). A Turing machine works with operations on symbols, not only mathematical symbols, but any kind of symbols, even mental symbols. A computer is the implementation of a universal Turing machine. This led to the development of programs that could generally solve (well-defined) problems (Gardner, 1985). Human intelligence, or at least a part of it, could be simulated on a computer. One of the earliest examples of these programs is the General Problem Solver by Newell, Shaw and Simon (1958). Because of its architecture and its foundation on theories of intelligence, the structure of this program was the beginning of a simulation for human cognition. The General Problem Solver, however, was limited in its possibilities: the only method in the program was the means-ends analysis and learning was not included.

In Human Problem Solving (HPS; 1972) Newell & Simon described their approach to human cognition as basically problem space/state oriented. Humans solve problems all the time. They do this by postulating a problem space, with initial, intermediate and goal states and with operations, of which a fixed set is combined in a method and in which search is conducted in a systematic way (Jorna, 1990). After the success of HPS Newell continued his approach in which he longed for a general theory of human cognition. Newell criticized the fragmented research after separated psychological phenomena in his article "You can't play 20 questions with nature and win" (1973) According to Newell it was necessary to focus on unified theories which included various aspects of cognition. Although there are many shortcomings in all parts, it is necessary from the beginning to look for an overall structure. In line with the (physical) symbol systems approach that he developed with Simon, Newell elaborated the concept of a production system as a description of human cognition. A production system is a scheme that specifies processes of information. It consists of a set of productions (condition-action rules) and a collection of data structures, i.e., expressions that encode information on which the production system operates. This basic structure is applicable to all architectures of cognition. This theory is simulated on a computer within an architecture for cognition. Structural requirements for human cognition are implemented in such an architecture. Components of this architecture are several memory structures, one or more central processors, representations as combinations of symbol structures and operations on these symbol structures. A very important assumption is that the description of these components is at a functional level. From 1975 until 1990 Newell's research resulted in many cognitive architectures. Newell's own results were published in *Unified Theories of Cognition* (1990). The resulting cognitive architecture is SOAR (States Operators And Results/Reasoning (van den Broek 2001)). Other cognitive researchers also developed and implemented architectures for cognition. More than 26 serious architectures were counted during the nineties of which the most important ones are besides SOAR (Newell, 1990), EPIC (Meyer & Kieras, 1997), 3CAPS (Just & Carpenter, 1992) and ACT-R (Anderson, 1993). The main goal of all these theories is the interaction between on the one hand a theoretical outline of human cognition on the basis of psychology and the neuro-sciences and on the other hand a practical simulation of these theories. The simulations often resulted in anomalies and shortcomings which lead to adjusted theories. At this moment ACT-R (Anderson, 1993) is more successful than SOAR, because of the fact that ACT-R tries to incorporate the functional as well as the intentional level and because learning in the sense of adaptation is accounted for. ACT-R is used in various kinds of practical applications, such as doing arithmetic and testing pupils. The theoretical development of architectures of cognition, for example ACT-R, is at the moment such that they can be used to model our earlier mentioned TCE-COG actor, not only at an intentional level, but also at the functional level.

With ACT-R we will design a representational response function actor that exhibits the behavior of actors that are used within TCE. In this way we can re-simulate the economic behavior of actors that do transactions. In this way loyalty, trust, opportunism and learning are not merely relevant at the intentional level, but are also accounted for at the functional level. This is not to say that with the implementation of ACT-R as a realization of TCE-COG actors, the same outcomes will necessarily turn up as in the non-cognitive situation. That remains to be seen. If so, the better for TCE; if not, we have a better understanding of the economic behavior of actors. This means that we have a win-win-situation. Whatever game we play in simulating the economic behavior of actors, if we loose TCE can be corroborated more strongly and if we win TCE can be adjusted in which case it will turn up to be stronger. In order to understand what conditions are relevant for a cognitive implementation of the actor we will first explain the basic elements of ACT-R

5 ACT-R

Within ACT-R, as in any other artificial intelligent architecture, the system functions because goal (s) have to be achieved. Goals may be winning the most of ten games of rock-paper-scissors, or realizing a maximum profit in 20 transactions in a certain market. To achieve these goals, one or more resources or actions have to be available. One of these actions has to be chosen. The choice ACT-R makes is based on the facts it has acquired and on its recent experiences, for example with opponents, earlier choices and strategies. Most problems are too complex to be solved within just one step. Complicated problems can be solved by making little steps one by one, such that less and less complicated problems remain. Within the SOAR community this is called ‘universal sub-goaling’.

The makers of ACT-R try to achieve plausible intentional level behavior as a result of behavior at the adequate functional level. ACT-R starts with the functional level as can be seen in the following remark: “*ACT-R consists of a theory of the nature of human knowledge, a theory of how this knowledge is deployed, and a theory of how this knowledge is acquired.*” (Anderson & Lebière, 1998). The dynamics of knowledge is achieved by changes in parameters at a lower level of description. Within ACT-R this is called the sub-symbolic level. The parameters reflect the probable usefulness and familiarity of a piece of knowledge. A high familiarity of a piece of knowledge means that it is very well known, e.g. one’s own name, where a low familiarity refers to a not well known piece of knowledge, like the name of an author of a certain book. This latter type of familiarity sometimes results in a situation in which you know the name of the author, but cannot actually retrieve it at this moment.

ACT-R keeps track of elements at the functional and sub-symbolic level and not at the intentional level. However, changes at the functional level - as a ‘result’ of changes at the sub-symbolic level - can be seen in the behavior of the model, that is to say at the intentional level (e.g., an actor - the ACT-R simulation - cooperates more than in the past). It should be noted that changes at this level are not to be seen in terms of an increase in one behavioral parameter (e.g., ‘the desire for cooperation’ parameter value has changed). Changes in behavior, such as an increase or decrease of cooperation, are the result of changes at lower levels (i.e., the functional or sub-symbolic level).

Changes in the behavior of the system do not arise randomly; they are the result of a striving for better. The question is how this improvement is measured and how this is implemented in ACT-R. What adaptation mechanisms are used? ACT-R adapts according to the principle of rational analysis (Anderson, 1990). Rational analysis tell us that each component of the architecture, although limited in computational power, is under a constant strive to optimize with respect to the environment. If we turn back to our TCE discussion this means that the goal is to make as much profit in the market environment with its participants. Anderson inclines to an evolutionary argument. The way choice is implemented serves is a good example of this principle. Sometimes more than one strategy or memory element can be used in ACT-R. In these cases a choice has to be made. That option will be chosen that, based on previous experience, has the highest expected gain and by consequence the lowest expected cost with respect to the expected gain of the outcome. The expected gain is updated after a strategy (or memory element) is used, so every experience contributes to the future estimation of gain of the strategy (or memory element). Future behavior depends on experiences of the past.

Not only the architecture is optimized according to rational analysis, task knowledge and knowledge acquisition are shaped by learning in the same way. Learning not only accounts for acquiring knowledge. It also aims at finding a good representation. There is no one way to represent knowledge in order to find the optimal. Rationality within ACT-R means: optimal adaptation to the environment. This may even result in not using available knowledge because the costs of using this knowledge are too high. In this way ACT-R is able to explore in order to acquire new knowledge and also to take into account that knowledge may be inaccurate.

5.1 The symbolic level

Within ACT-R the memory component consists of two kinds of knowledge: declarative and procedural knowledge. Memory is a functional component of the actor (at a functional level of description), whereas the content of memory is symbolic in form. In ACT-R this is called the symbolic level.

Declarative knowledge sometimes is described as explicit knowledge and most of it consists of facts. It is knowledge we are aware of and which we can usually tell to another person. An example of declarative knowledge is “Amsterdam is the capital of the Netherlands.” In ACT-R the declarative knowledge consists of chunks (Miller, 1956; Servan-Schreiber, 1991). They may contain perceptions, facts, (the solution of) a previous goal or the current goal. Two examples of chunks are:

Oldgoal	Currentgoal
Isa multiplication	Isa division
Addend1 three	Num1 nine
Addend2 three	Num2 three
Answer nine	Answer Nil

‘Isa’, ‘Addend1', ‘Addend2' and ‘Answer’ are slots in the chunk ‘Oldgoal’. ‘Isa’ is a special type of slot that specifies the chunk. In this case the knowledge stored in the chunk concerns a multiplication. ‘Three’ and ‘Nine’ are references to other chunks in declarative memory (i.e., numbers). ’Nil’ in the ‘Answerslot’ of the ‘Currentgoal’ chunk refers to the fact that an answer has not yet been found for the current problem. The whole chunk ‘Oldgoal’ represents the fact that the multiplication of three with three equals nine (Anderson & Lebiere, 1998).

Procedural knowledge is also known as implicit knowledge. It is knowledge that we are not generally aware of. An example is the knowledge of how to ride a bike or how to rotate an object in mental imagery. The concept of a ‘strategy’ closely resembles procedural knowledge. ACT-R uses the same distinction, where procedural knowledge basically specifies how to bring declarative knowledge to bear in solving problems. Procedural knowledge is represented in production rules that operate on chunks of declarative knowledge. The rules can be used if the specific condition part is met. If the current goal is to find the answer of dividing nine by three, the rules concerning divisions can be used. If the problem is somewhat more complex, a production rule can transform the current problem into another problem. For example, if we can't find the answer to our current division-problem, we will try whether we can find the answer by rewriting our problem to an multiplication-problem. In production rules not only mental actions can be the result, but also motor actions.

Declarative memory acts both as factual and as working memory. Procedural knowledge is stored in procedural memory. The information flow within ACT-R is given in figure 6.

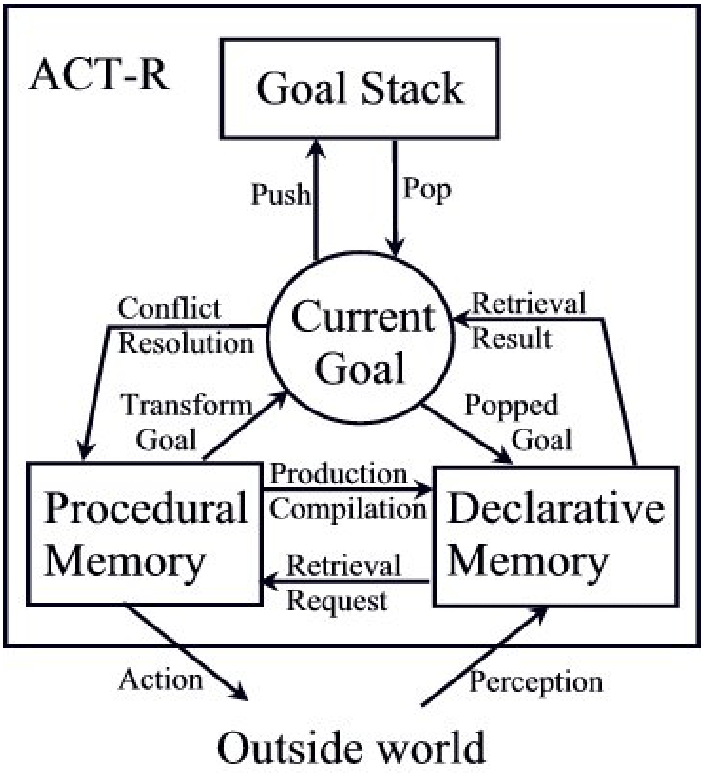

Figure 6: The information flow in ACT-R

The current goal and the goal stack determine the way information is processed. We will discuss this later. Next to the symbolic level - where representations play an important role and which entails the recognize act cycle, typical for production systems - ACT-R has a sub-symbolic level. This sub-symbolic level refers to the neurological substrate of every (natural) functional architecture.

5.2 ACT-R's sub-symbolic level

ACT-R's sub-symbolic level consists of parameters that qualify the value of a certain piece of knowledge. Every piece of declarative knowledge (i.e., every chunk) has a parameter that tells us how familiar it is. This familiarity or base-level activation depends on how often and how long ago a chunk was used. Next to the base-level activation there is context association. Context association refers to the fact that although, for example you cannot immediately remember the name of an author, you will know the name when you hear the name of a book she has written. Depending on the current goal/context, the context association is the bonus the chunk gets.

Every piece of procedural knowledge also has parameters that give a qualification of its adequacy. Based on these parameters a choice can be made between two production rules.

Together these parameters form estimates of the expected gain of the production rule. The parameters are concerned with the chance of reaching the goal and the costs of reaching that goal. If we have two different ways to find an answer, for example, to a subtraction-problem (five minus three) - namely to find a corresponding addition problem (two plus three equals five) or to subtract one by one by mentally raising or bringing down fingers (five minus three equals four minus two, which equals three minus one, which equals two) - that strategy will be used that takes the least effort (least cost) and has the highest chance of finding the (right) answer. It should be noted that every experience modifies the underlying parameters of the knowledge used, both procedural and declarative. If ACT-R starts to learn math, it will start by using the 'counting on fingers' strategy. Because ACT-R is not yet confident of the answer of the corresponding addition-problem - because it has done the addition only once - it may not yet have found the answer to the corresponding addition problem at all. When the answer is reached - and feedback is given -, the answer is added as a chunk to the declarative memory. When the problem emerges again, it can directly be retrieved from memory or, if memory fails, it can be calculated again and ACT-R's memory is strengthened. Apart from the addition of a chunk to the declarative memory, the parameters of the used production rule are adjusted in the light of this recent experience. Both methods - chunking and adjusting - are part of the mechanisms in ACT-R concerned with learning.

5.3 Learning

Learning takes place in declarative as well as in procedural memory and at the symbolic as well as at the sub symbolic level.

Symbolic learning: When a problem is solved, new chunks are added to declarative memory at the symbolic level. Besides the possibility that a chunk is created internally by the goal processing of ACT-R itself, a new chunk in declarative memory can be a perceptual object. ACT-R's internally created chunks are always old goals. Learning new production rules is a more intricate process. Production rules are learned from examples. When enough examples have passed, ACT-R is able to generalize them into a new production-rule. The discussed production-rule for solving a subtraction by finding the corresponding addition can be deduced by ACT-R. When deducting this rule, the sub-symbolic parameters are also estimated.

Sub-symbolic learning: At the sub-symbolic level parameters are adjusted when rules are applied or chunks are used. The learning of sub-symbolic parameters is guided by Bayes' Theorem (Berger, 1985). It works in a selective way according to the principle of rational analysis. The base-level activation of a chunk reflects the probability that it is needed. The probability that a chunk will be needed is relatively high, if it was retrieved a number of times in the immediate past. If a chunk has not been retrieved for a long time, the probability that it will be needed now is only small. Each time a chunk is retrieved, it's base-level activation increases. Each time it is not used, it decreases (every second it is not used it decreases, so the memory constantly degrades). The other parameters, like the expected chance that a production rule will lead to success, are estimated in a similar way. Parameters increase each time it successfully leads to a goal, and decreases each time the goal is not reached. Due to this principle of rational analysis the adjustment of the parameters correlates with the usability of the rules and chunks. Like an evolutionary process they will select themselves. To improve this selection mechanism, a little noise/chaos is added at the sub-symbolic level. From the above description it is not yet clear how information is processed. We will discuss this in the next section.

5.4 The processing of information.

The flow and processing of information in ACT-R is determined by the current goal and the goal stack. Most problems in real life - and in ACT-R - are too complex to be solved in just one step, so they are solved in parts. For example, to solve a ‘get cash’ problem, two sub-goals - find a cash-dispenser and withdraw money from the dispenser - are made. When these sub-goals are solved, the problem is solved. The sub-goal ‘find a cash-dispenser’ can be hard to solve and may therefore have to be broken up into sub-goals, such as ‘ask for directions’ and ‘walk according to the instructions into the directions’.

In ACT-R a production rule creates one or more sub-goals out of the main goal. This sub-goal can be solved more easily. The main goal is put away on a stack. ACT-R now tries to solve the sub-goal and in doing so it possibly solves the main goal. Within ACT-R putting another goal in the current goal position is called pushing a new goal. This sub-goal can also lead to a further breaking up of the problem into easier solvable pieces. This sub-goal is also put on the goal stack and a new sub-goal takes its place in the current goal position. This happens until the problem cannot be divided up anymore by ACT-R or until a sub-goal is solved. The sub-goal is then removed from the current goal position (i.e., “popped” in ACT-R language). The last not (yet) solved sub-goal then becomes the current goal. This means that the goal stack is ordered as ‘last in, first out’. In this way the goal problem gets solved by first solving all small parts. The production rules applicable in the given problem situation guide the requests from the declarative memory. The transformed/newly created sub-goals will also trigger other production rules (designed for solving other problems) to become active. This concludes the general overview of the information processing of ACT-R.

We now have some basic understanding of the architecture and procedures of ACT-R and of how it behaves. Earlier we made a distinction in various kinds of actors and we stated that a plausible cognitive actor should be the instantiation of a ‘representational response function system’ (RRF-system). Does ACT-R meets the demands that a RRF-system poses to an actor and what does this mean for our argument with regard to TCE and its lack of cognitive plausibility?

6 The ACT-R architecture as a RRF-system and as a cognitive plausible TCE-actor.

As stated before, the RRF-system demands form the bridge between the intentional level of the TCE-actor and the functional level of an Architecture for Cognition (AFC). At the intentional level we want the actor to be able to lie, to cheat, to be opportunistic, to be trusted and to trust. He (or she) must be able to develop a trust based relationship that may be broken in order to fulfil (selfish) goals of making more profit. In order to do so, the actor must be capable of getting information, of negotiating, of giving (mis)information and of deliberating on possible actions to be undertaken. In the course of time this actor will learn what kind of behavior best suites its goals and will consequently adapt his behavior. Optimization takes place with respect to the goal of the actor and on the basis of the bounded rational faculties.

The functional demands based on the RRF-system will have to be accounted for by the functional model underlying ACT-R. We will now determine whether ACT-R fits our RRF-demands by examining them one by one: *perception*, *interaction* (including learning in the sense of habit formation), *representation and interpretation* (including learning in the sense of chunking and creation) and *autonomy and self-consciousness* .

Perception: An actor has sensors with which he can acquire information from the environment. This can be realized on purpose by means of effectors, by which an actor can influence or perform actions with regard to his environment. The actor is part of the environment, meaning that he can manipulate himself. In ACT-R executing procedural knowledge can result in effecting the environment or changing internal goals. Declarative knowledge can be the result of perception. Self-manipulation in the sense of learning is done automatically.

Interaction: Various components can interact leading to new functions. The combination of a

sensor, a memory, a processor and an effector leads to a communication channel. This is also the case in ACT-R. In this way a combination of functional elements results in a higher function.
Representation and interpretation: An actor has one (or more) information processors. The specification of how goals are pushed and popped and how a production rule fires and requests information from the declarative memory constitutes the information processor(s). An actor also has a memory and the content of memory is representationally not well-defined and dynamic. In ACT-R the memory consists of two types of knowledge: the declarative and procedural, but they can stand for anything. The content of memory is different from the processor. The content is symbolic in many ways: propositions, icons, images, episodes, scripts or procedures. This implies that the content is representational. In ACT-R memory can be symbolic and sub-symbolic, procedural and declarative. This differs from the way knowledge can be processed. An actor has personal goals that are present in the form of representations. An actor strives toward the realization of these goals. An example is the maximization of profit within the TCE model. An ACT-R actor also has goals and tries to optimize its behavior with respect to its goals and environment. This optimization in ACT-R is called rational analysis and is one of the key elements of the ACT-R theory.

It can bee seen that ACT-R is very well suited to devising our RRF-system actor. The one thing ACT-R actors do not yet seem to have is awareness. We would like to call it representation of representations. However, this very complicated concept, is not necessary in solving most of the problems encountered in the (economic) world.

Although ACT-R fulfils all but one of the requirements, it still may be difficult to see how with ACT-R a cognitive plausible TCE-actor (TCE-COG actor) could be developed that can lie, cheat, be opportunistic, be trusted and develop trust. We will exemplify our conjecture that with ACT-R adequately behaving TCE-COG actors can be devised.

Suppose that within a TCE perspective an ACT-R actor will have the goal to get as much profit as possible in a market consisting of six buyers and six suppliers (Klos, 2000; Helmhout, 2001). This means that the buyer actor has to make choices with which supplier actor he will perform a transaction in order to get as much profit as possible. The supplier actor also tries to get as much profit and will try to make transactions with the most profitable buyers. Suppose several rounds are played and that in each round each buyer sends messages to three suppliers and each supplier can supply to a maximum of three buyers. In this example the market will be the organizational mechanism that matches the buyers with the suppliers. Some buyers will get lucky to meet a supplier that wants to supply for them, some will be unattractive for the suppliers and will get no supplies. They will then produce for themselves in order to make some profit. In the same way some suppliers will get no orders, but they won't buy from themselves and will not make any profit. Both, suppliers and buyers, can decline an offer to cooperate and by doing so they may harm the other. In that case an actor has chosen to invest in the wrong partner and is not able to make profit that round. The actors will have 25 rounds to perform transactions and gain as much profit as possible. In order to see more clearly the effects of learning, the actors will go trough many of such sets of 25 transactions, each time in a changed, and therefore new, market. As said earlier (section 5), ACT-R optimizes with respect to its past. This means that each time a strategy (or memory element) is used, an evaluation takes place afterwards (both declarative and procedural memory are evaluated). The most simple actors to be devised in order to perform on this task will have three procedural pieces of knowledge. One will be called 'cooperate', the second 'switch' and the third 'produce for oneself' . In 'cooperate' the same supplier as in the last round is chosen, in 'switch' a more attractive other supplier is preferred. Producing for oneself is evident. The procedures will give an estimate of their expected gain with respect to the environment and the goal ('making profit'). Next to these strategies partner specific information is needed. This piece of declarative knowledge that will be called 'partner-info' reflects previous experiences of this actor with a specific partner. The three pieces of procedural knowledge will both make use of these pieces of declarative knowledge. By going through 25 rounds (transactions), the actors will have adapted their sub-symbolic parameters reflecting usefulness and possible gain of the pieces of knowledge, i.e., the usefulness of the 'cooperate', 'switch' and 'produce for oneself' rule. As a result a score will be associated with every possible partner.

The 'switch' strategy may give profit in the short term. However, because of switching, the partner may be harmed and this partner will give the defector a low score making it very implausible that he will be accepted again for a transaction. Having only six possible partners it is possible that all have been offended once and so no one is willing to perform a transaction, anymore. Such an outcome depends on the behavior of the other players in the market. When they are more reliable they may be preferred, even when a successful transaction with the one time defector will generate more profit. Not only one's own behavior will guide one's own choices, the behavior of the partners will also guide the choice. When the partner in the last round has run away to another partner, it is very unlikely that the next choice will be made in favor of the 'cooperate' strategy. Then another partner will be chosen. The scores with respect to the possible partners are the result of the realized profit, which depends on the trustworthiness and the profitability. In a situation where a buyer chooses to continue a relation with a less profitable partner, trust can be said to have been evolved. After a couple of rounds, the actor will have adapted its weights depending on experiences. When he is solely confronted with very opportunistic partners, he will be likely to have developed a high score for the 'switch' or 'produce for oneself' strategy. When the actor has experience with markets that are divided into trustful and opportunistic partners, the 'cooperate' strategy can be rated high. This is reasonable because a trustful partner is very valuable and only a recent defect of the partner will make a switch to a possible opportunistic partner valuable. It can also be the case that the actor has become opportunistic (high switch score) himself and that he seizes a opportunity when he gets one because he cannot trust the other. It would be interesting to investigate what would happen to an actor that has a history of very opportunistic markets and is placed in a new market consisting of actors having a very loyal market history. Will the opportunist change the loyal players, or can the loyal players make the opportunist loyal?

7. Conclusion: the march toward a TCE-COG actor implementation

In discussing TCE we showed that TCE uses assumptions about the architecture of actors that do not fit with what is known from evidence on cognition and intelligence in cognitive science. Although TCE works with trust, loyalty, opportunism and adapted behavior of actors, nothing is mentioned that supports the so called intelligence of actors. Bounded rationality is assumed but not incorporated, learning is assumed but not accounted for. This is not a reason to dismiss TCE. We argued that what might be called a flaw of TCE is a consequence of a division of scientific labor. It is not the aim of economic theory to come up with the foundations of cognitive science. However, we do want TCE-actors to behave intelligently, that is to say in a rational way. TCE-actors *are* or *should* be cognitive TCE-actors. If they *are* cognitive actors, then we should be able to discern a cognitive architecture and mental representations. We tried to find or reconstruct these features, but we were not able to do that. If they *should* be cognitive actors they have to be extended with an architecture and representations. That is the proposal we did in the foregoing sections.

In order to determine what boundary conditions are relevant for plausible cognitive actors we analyzed which cognitive components in general can be discerned. We distinguished perception, interaction, learning, representations and self-representation. Based on these actor components we argued that a minimal cognitive actor should consist of perception, interaction, learning and representations. We called this kind of actor a representational response function system, a RRF-actor. In the study of cognitive architectures and of artificial intelligence in general behavior of actors can be described at various levels of description. We demonstrated that TCE describes actors at an intentional level of description, taking for granted that at lower levels of description the components and elements are adequate. We argued that this adequacy can be questioned or even stronger is not accounted for within TCE. In architectures of cognition one basic rule is that the physiological, functional and intentional levels of description should be interconnected and if one level of description is left out it should not be because of incompleteness, but for reasons of load of explanation. A cognitive architecture like Soar leaves out the physiological/physical level and ACT-R mimics the physiological level in simulations in its learning/adaptation mechanisms. In TCE and its actor implementation the behavior of buyers and suppliers is only in terms of the intentional level of description. Trust, loyalty, opportunism and learning are only accounted for at the intentional level of description. The mechanisms behind or on the basis of this behavior are not accounted for.

In a nutshell our approach is to complete the behavior of TCE-actors at the intentional level by equipping them with cognitive mechanisms at the functional level. Out of the large amount of architectures of cognition we choose tho use Anderson's ACT-R architecture (Adaptive Control of Thought or Atomic Components of Thought). We gave four reasons. First, ACT-R can be used at the intentional and the functional level of description. Second, ACT-R tries to account for the physiological level, because it simulates neurological patterns. Third, ACT-R is the far most developed implementation of cognitive components. Fourth, ACT-R does account for learning in terms of adaptation mechanisms.

Our approach is not to decline TCE or other economic theories, on the contrary we take what they state for granted at the intentional level, but try to find the cognitive underpinnings and if these underpinnings are missing we suggest architectures that can fill in the gaps. ACT-R is such a completion.

In the discussion of TCE and in extending the work of van den Broek (2001), Klos (2000) and Helmhout (2001), two research lines can be followed. The first line is the implementation of cognitive architectures as part of TCE-actors such that TCE-COG actors are realized. The second is the implementation of various kinds of governance structures. Concerning the implementation of cognitive actors trust, loyalty and opportunism have to be implemented in an ACT-R simulation. Concerning trust we take the perspective that trust is based on reducing uncertainties and working with regularities. It is a mental action consisting of estimations concerning regular behavior of the actor itself, the behavior of others and profit in the future. Trust changes because of a decay of the actor's memory components, of experiences with one's own behavior and experiences with the behavior of others. Trust, therefore, is representation and adaptation based. It requires memories and learning mechanisms. Besides the adaptation of trust there also is what can be called base trust. No intelligent actor can live without taking some things for granted. If the base trust is high we could call it loyal behavior, providing that no sudden ruptures appear in the relations the actor entertain. If the base trust is low, we could call it opportunistic behavior and perhaps this low base trust is very little affected by sudden ruptures in existing relations. The change in trust or the level of dynamics of trust is something that can be experimented with in ACT-R simulations models.

The second line of reasoning concerns the various kinds of governance structures. Klos (2000) worked with markets in which buyers and suppliers negotiate and form and break relations. Other coordination mechanisms can also be investigated. One could think about hierarchies, bureaucracies, standardized structures or meta-plans. The cognitive line and the governance line are interconnected but parallel. Because we think that cognitively plausible actors are more interesting we start there. However, it is only the beginning. Once a minimal TCE-COG actor simulation is working - and that is the first that will be done - it is almost necessary to further develop both lines of research: the extend cognitive and the governance variation.

From the two lines of reasoning it can be concluded that we have more pretensions concerning TCE then concerning ACT-R. Although we believe that social elaborations of single intelligent architectures are relevant for cognitive science, our main focus is the cognitive plausibility of economic and social actors.

References

Anderson, J.R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J.R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates

Berger, J.O. (1985). *Statistical decision theory and Bayesian analysis*. New York: Springer Press.

Broek, van den, H. (2001). *On agent cooperation: the relevance of cognitive plausibility for multi-agent simulation models of organization*. Groningen: SOM-Thesis, University of Groningen.

Coase, R.H. (1937). *The nature of the firm*. Economica NS, 4 (16), p. 386-405.

Dennett, D. C. (1978). *Brainstorms: Philosophical essays on mind and psychology*. Hassocks, Sussex: Harvester Press.

Dennett, D. C. (1987). *The intentional stance*. Cambridge (Mass): The MIT Press.

Dennett, D.C. (1991). *Consciousness explained*. London: Allan Lane, The Penguin Press.

Dennett, D.C. (1998). *Brainchildren: essays on designing minds*. Cambridge (Mass): The MIT Press

Dietrich, M.. (1994). *Transaction Cost Economics and beyond, towards a new understanding of the firm*. London: Routledge.

Fodor, J. A. (1975). *The language of thought*. New York: Thomas Y. Crowell.

Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.

Gazendam, H. W. M. (1993). *Variety controls variety, on the use of organization theories in information management*. Groningen: Wolters-Noordhoff.

Gazendam, H.W.M. & Jorna, R.J. (1998). *Theories about Architecture and Performance of Multi-agent Systems*. University of Groningen: SOM Research Report 98 A02.

Goodman, N. (1968/1981). *Languages of Art*. Brighton, Sussex: The Harvester Press. (Second Edition)

Helmhout, J.M. (2001). *Make or Buy: decisions based on TCE. Groningen*. Master Thesis Fac. of Management & Organization, University of Groningen (in Dutch).

Jorna, R. J. (1990). *Knowledge representation and symbols in the mind*. Probleme der Semiotik. Tübingen: Stauffenburg Verlag.

Jorna, R.J. (1998). Knowledge Management really is about Representations. *The Semiotic Review of Books, Vol 9.2* , pp. 5-8.

Jorna, R.J. (2000). *De 'zwarte doos' in de bedrijfskunde: cognitie in actie*. (Inaugural lecture; June 2000; University of Groningen; Faculty of Management & Organization. (The "black box" in management science: cognition in action)

Just, M.A. & Carpenter, P. N. (1992). A capacity theory of comprehension: Individual

differences in working memory. *Psychological Review, 99*, p.122-149.
Klos, T.B. (2000). *Agent-based Computational Transaction Cost Economics.* Labyrint publications, Capelle aan de IJssel.
Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
Meyer, D. E. & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part1. Basic mechanisms. *Psychological Review, 104*, p. 3-65.
Michon, J.A., Jackson, J.L. & Jorna, R.J. (2002). Semiotic Aspects of Psychology. In: R.Posner, K. Robbering and T.A. Sebeok (Ed.). Semiotics: *A Handbook on the Sign-Theoretic Foundations of Nature and Culture*. Berlin: de Gruyter. (In press)
Miller, G.A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*, p. 81-97.
Millican, P. J. R. & Clark, A. (1996). *The legacy of Allan Turing*. Oxford: Clarendon Press.
Newell, A. Shaw, J.C. & Simon, H.A. (1958). Elements of a theory of human problem solving. *Psychological Review, 65*, 151-166.
Newell, A. (1973). You can't play 20 questions with nature and win. In W.G. Chase (ed.). *Visual Information Processing*. New York: Academic Press.
Newell, A. (1982). The knowledge level. *Artificial Intelligence, 18*, p. 81-132.
Newell, A. (1990). *Unified theories of cognition*. Cambridge (Mass): Harvard University Press.
Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
Newell A. & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (ed.). *Cognitive skills and their acquisition*. Hillsdale, NJ: Lawrence Erlbaum.
Noorderhaven, N.G. (1993). *The argumentational texture of transaction cost economics, Department of economics research memorandum*. KUB: Tilburg.
Nöth, W. (2000). *Handbuch der Semiotik (2e Auflage)*. Stuttgart: Metzler Verlag.
Numan, J.H. (1998). *Knowledge based Systems and Companions*. Capelle aan de IJssel Labyrinth Publications
Posner, M. I. (ed.). (1989). *Foundations of cognitive science*. Boston (Mass): MIT Press.
Pylyshyn, Z. W. (1984). *Computation and cognition: Towards a foundation for cognitive science*. Cambridge (Mass): MIT Press.
Servan-Schreiber, E. (1991). *The competitive chunking theory: Models of perception, learning, and memory*. Carnegie Mellon University: Pittsburg, PA (Unpublished doctoral dissertation).
Simon, H.A. (1947/1976). *Administrative behavior*. New York: Macmillan.
Simon, H.A. (1960). *The new science of management decision*. New York: Harper & Row.
Simon, H. A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.
Williamson, O. E. (1985). *The economic institutions of capitalism: firms, markets, relational contracting*. London: Collier Macmillan.

Groningen, April 15, 2002

Correspondence address: r.j.j.m.jorna@bdk.rug.nl
René Jorna (& Luit Gazendam)
Knowledge Management
Faculty of Management and Organization
University of Groningen,
P.O. Box 800, 9700 Av Groningen
The Netherlands