In *Cognition, Culture & Design*, S. Bagniara, S. Pozzi, A. Rizzo & P. Wright (Eds.), Sienna: Instututo di Sienze et Tecnologie dela Cognizione, 2002, pp. 137-142

Socio-technical Systems Analysis: Which approach should be followed?

Dimitris Nathanael **Nicolas Marmaras** **Bill Papantoniou** **Nikos Zarboutis**

ErgoU, Dept. of Mechanical Engineering
National Technical University of Athens,
9 Iroon Polytechniou st., GR-157 80, Zografos, Greece

+30 10 7723938 dnathan@central.ntua.gr
+30 10 7723492 marmaras@central.ntua.gr
+30 10 7723938 billpapa2001@yahoo.com
+30 10 7723938 nizar@central.ntua.gr

ABSTRACT

The aim of the present paper is to contribute to the debate in the cognitive ergonomics community as for the adequacy of different approaches for the analysis of socio-technical systems in view of designing artifacts facilitating human work. Using a taxonomy of socio-technical systems as frame of reference, the theoretical origins and the main application domains of three well-known approaches are examined. This permits to extract useful conclusions as to their appropriateness and effectiveness for the ergonomic analysis and design.

Keywords

Socio-technical systems, Cognitive Systems Engineering, Distributed Cognition, Activity Theory, Design.

INTRODUCTION

Cognitive ergonomists intervene in socio-technical systems with the aim to design (or contribute to the design of) artifacts facilitating human work. Examples of such artifacts are information technology systems and organizational settings. The particularity of Cognitive Ergonomics is the use of the joint human-technology system as frame of reference (Hollangel, 2001). Successful interventions depend heavily on the system's analysis carried out, as it indicates challenges for the type of intervention and provides the frame for predicting the effects of the alternative design solutions.

In recent years a number of approaches are proposed for the analysis of socio-technical systems and the ergonomic design of artifacts that become part of joint cognitive systems. Among the most well-known ones are Rasmussen's Cognitive Systems Engineering (Rasmussen et all. 1994), Hutchins' Distributed Cognition (Hutchins 1995) and approaches based on Activity Theory (Leontiev 1974, Engeström 1997).

These approaches have been used in a wide range of work domains with different degrees of effectiveness regarding the ergonomic design. This remark raises the question as to their appropriateness and usefulness for any work domain to which cognitive ergonomists intervene.

In the present paper we examine the theoretical origins of the above approaches as well as their main application domains. To do this, we use a taxonomy of socio-technical systems as frame of reference. The taxonomy concerns the nature of the various socio-technical systems and is based on the main forces governing their behavior.

The confrontation of the examined approaches with the nature of different socio-technical systems, permits to extract useful conclusions as to their appropriateness and effectiveness for the ergonomic analysis and design.

TAXONOMY OF SOCIO-TECHNICAL SYSTEMS

Socio-technical systems could be positioned in a continuum delimited by two poles:

Pole A

Pole A attracts systems governed by well-determined laws that cannot be altered within the life span of the system. Examples of such systems are industrial process plants, transportation vehicles and data mining systems. The laws governing the above systems are universal in the sense that they involve basic energy, mass and information transformations. They are manifested by a stable transformation function that predominates over the system, i.e. it induces a stable set of constraints that largely determine the system's behavior.

Human agents in such systems recognize the constraints generated by the transformation function and their activity is tuned by them.

When a trained operator is found in such a highly constrained work setting, such as the control room of a nuclear power plant, he faces a causal world where the physical process itself is both the object and a major part of the work environment. In this sense the causality of the physical process provides the operator with both a goal to achieve and the corresponding rules to obey in order to achieve this goal. The water – steam temperature and pressure relations of the reactor cooling plant are an integral part of the process of transforming mass to energy, so achieving a balance between them is a goal towards the objective. At the same time, the relations between them provide the operator with a set of stable constraints that greatly restrain alternative courses of action.

As a result, there is little opportunity for other constituent elements of the system to emerge as determinants of its overall behavior (e.g., society, finance). Because the transformation function largely defines the objectives of such systems, they can be modeled as causal without losing much of their actual behavior characteristics.

Pole B

Pole B attracts systems in which the transformation function is not stable, i.e., the laws governing it are vague and/or may change considerably within the life span of the system. Typical examples of Pole-B-attracted systems are banking institutions, schools or machine workshops.

In the absence of universal stable laws, other elements, e.g., humans by means of their intentions, predominate as determinants of system's behavior. There is no intrinsic causality that dominates over the system, since a number of parallel objectives may be at play at any moment. Humans utilize the vagueness and unstableness of the constraints that characterize their transformation function. They exploit the many degrees of freedom for action depending on their personal intentions, short-term goals, history or situational constraints. Additionally the intrusion of human intentionality and artifacts shatters the homogeneity of temporal-spatial scale, (Latour, 1987; Lemke 2000); the history of the system and/or its components may spread in a different timescale.

The description between Pole A vs. Pole B shares common ground with the distinction between "tightly vs. loosely coupled systems" proposed by H.A. Simon (1976), Perrow (1999) and Checkland's (1981) "hard vs. soft systems". Simon observed that in loosely coupled systems, goals and intended outcomes are often broad, ill defined and vague. According to Simon, in such systems, street-level implementers typically have no objective ways to tie individual actions to broad end goals. Perrow adopting a functional view, defines coupling "as the amount of 'slack, buffer or give between two items". According to Perrow "loosely coupled systems are characterized by decentralized operations, mission orders, ambiguous performance standards and flexible control mechanisms". A similar distinction was made by Peter Checkland in the seventies, who tried to develop methodologies across the spectrum from: "the relatively 'hard' systems involving industrial plants characterized by easy to define objectives, clearly defined decision-taking procedures and quantitative measures of performance… (to) …'soft' systems in which objectives are hard to define, decision-taking is uncertain, measures of performance are at best qualitative and human behavior is irrational" (Checkland, 1981).

We support that the effectiveness of the approaches examined hereafter depends on the positioning of the system they are applied to, on the continuum defined by Poles A and B.

THE COGNITIVE SYSTEMS ENGINEERING APPROACH

In Rasmussen's Cognitive Systems Engineering, the work domain is modeled by a means ends abstraction hierarchy. The objectives of the work domain studied are represented by a goal structure that must be well defined or as the author puts it "In general, objectives and goals must be defined for each level in the means – ends space and the objective of a function at one level is found at the next higher level". (Rasmussen et al. 1994, p. 42)

This goal structure follows the basic levels of abstraction with goals belonging to higher levels of abstraction clearly predominating over lower level ones. Actually, what the abstraction hierarchy maps are the constraints that govern the system's behaviour from general (higher levels) to very specific (lower levels). Constraints allocated to higher levels tend to originate from the objectives of the system and the underlying social intentions. This assumption works well when the purpose of the system is clear and the relevant goal structure is stable and thus identifiable to a large extent. This is usually the case of systems that are primarily governed (i.e. constrained) by laws of nature like a chemical process plant or air traffic control.

In this approach the human is seen as a system element that exhibits more or less rational behaviour vis-à-vis the system's objectives. The individual actor's intentions are well constrained by the purpose of the system and the (usually) tightly coupled nature of the process. For this reason their influence on the structure of the work domain may be considered negligible for the purposes of the analysis. As Rasmussen phrased it "the task of the operating staff is basically to ensure that the functioning of the automatic control system actually reflects the intentionality of the original design while their personal goals and preferences have little significance" (ibid., p. 50)

The approach was developed in view of designing control settings or software tools for systems that exhibited the above characteristics. For example, in addition to industrial process control, the abstraction hierarchy has been used to represent complex work domains such as large sea vessel manoeuvring (Petersen & Nielsen, 2001), patient physiological monitoring in intensive care (Sharp & Helmicki, 1996) and Software debugging (Yoon & Garcia, 1996). These domains share common characteristics such as unambiguous goal structure, well defined success criteria, highly skilled operators and at least for the two first examples a workflow with rigid time constraints. For all of the above work domains, it is possible to identify a more or less stable transformation function and thus a hierarchical representation of the work domain is both feasible and rich in content, resulting in a useful approximation of reality.

The authors (Rasmussen et al. 1994) have also expanded the Abstraction Hierarchy framework to include systems that lie close to pole B. As discussed in the introduction, these systems tend to have an ambiguous and dynamic goal structure much influenced by different forms of human intentionality (laws, policy, regulations).

In Rasmussen's words "….many degrees of freedom remain to be resolved by situational and subjective criteria by the staff at the intermediate levels of an organisation. This in turn implies that the individual actor faces a work environment in which the regularity to a considerable degree depends on the intentionality brought bear by colleagues. The intentional structure is, therefore much more complex and dynamic than for the industrial process systems" (Rasmussen et al. 1994, p. 52)

In order to represent the above complex intentional structure at the means – ends, part - whole space the authors (Rasmussen et al. 1994) make use of the concept of views. This results in a 'pseudo- hierarchy' of goals. The concept of views inevitably signifies possibility of conflicts - contradictions between goals since different views have independent goal structures.

As a result of the above it becomes extremely difficult to identify a stable transformation function that effectively bounds human actors' behaviour. Observations of real work settings will only reveal a particular instance of this productive process and any inference on the underlying causality will be speculative and not reliable for prediction purposes.

THE DISTRIBUTED COGNITION APPROACH

Hutchins' Distributed Cognition (DC) focuses on the organization of teamwork and specifically on the propagation and processing of information through a system comprising people and artifacts through representational states of mediating structures. These mediating structures include internal as well as external knowledge representations (knowledge, skills, tools etc.) (Hutchins, 1995).

The primary unit of analysis in Distributed Cognition is a "culturally constituted functional group" or a socio-technical system.

DC recognizes a relatively persistent structure to the socio technical system under study. That is it tries to identify a more or less constant grid constituted by physical – computational but also cultural & cognitive resources and constraints, where the distributed cognitive activity takes place and is left free to take a set of possible trajectories – pathways bounded by the above structure.

Although, Distributed Cognition acknowledges system intentionality as socially and culturally determined and gives great attention in the systems' historical development and evolution, in its typical high temporal resolution analyses it adopts a pragmatic approach. It considers a clear objective for the system and a goal hierarchy that is to a large extent identifiable (i.e. a stable transformation function is defined).

These assumptions work well in highly constrained work settings where i) the coupling of human agents with the environment is mainly driven by stable constrains and ii) the objectives of the system have a causal structure which is reflected in these constraints.

Indeed, the prototypical fields of development and application of Distributed Cognition (bridge of a navy ship, airplane cockpit) were highly rationalized domains (Hutchins & Klausen, 1996). In such systems the persistent structure is rich and not very deep, so the set of alternative trajectories of the distributed cognitive activity can be identified. In an airplane cockpit or on the bridge of a ship, the constraints imposed by a well defined objective obeying universal laws, the physical layout of the workspace, the affordances of the various artifacts, the stabilized norms and rituals of the trade, as well as the sequential nature of the activity offer a rich set of tangible boundaries and effectively delimit the number of alternative courses of action. Individual human agents have little "freedom" to influence the overall causality by means of their personal intentions. Individual agents may be responsible, and therefore alter, a shallow section of this goal structure, but the system is typically producing an overall behavior without there being a representation of the overall pattern anywhere in the system (Hutchins, 1995).

As discussed in the introduction, in domains close to pole B, the collaboration between agents may be based more on contradictions than on a dominant goal structure and constraints imposed by the representational media. The system's behavior is mainly governed by social conventions that are both vague and subject to change. Therefore, in an analysis mainly based on observable evidence, it may be extremely difficult to frame the system sufficiently in order to get useful results of its behavior.

In an attempt to apply Distributed Cognition in domains close to pole B, the term "activity landscape" has been proposed (Kirsh, 2001) to denote the structured environment in which specific work tasks unfold. Defining such an activity landscape is an extremely difficult task. Kirsh (2001) notes, that in domains such as office work, there is a need for a deep abstraction in order to arrive in anything tangible – stable. The same author points out that one of the main difficulties in defining the activity landscape of such systems is the inability to identify a stable goal structure.

Distributed cognition offers a valuable set of concepts to analyze cognitive work, which is well adapted for work systems characterized by multiple agents and a rather stable transformation function. Research is being carried out for the application of these concepts in less structured work domains. However, in such cases the analyst may not be able to fully exploit the main strong point of the approach, namely "the external representational state propagation" which is directly observable in highly distributed systems.

APPROACHES BASED ON ACTIVITY THEORY

Activity Theory (AT) was developed to study man as an actor in a cultural-historical context. It is based on the idea of the dual process of man and artifacts shaping and being shaped by social and physical environment. AT places emphasis on human activity with its focus on issues such as consciousness, motivation and history.

According to AT the objectives of the work domain are represented by the "object" of Activity. The "object is both the motivation of the activity and its projected outcome. As such it determines the horizon of possible actions for the activity but can only be revealed at the process of doing (Kuutti, 1996). The theory's focus on the actor's goals and intentions through the powerful notion of 'object' makes it a useful descriptive tool in the case of systems close to Pole B. In such systems the absence of a rigid transformation function leaves room for the actors' intentionality to come forth, making imperative to consider this intentionality during the analysis of the system. Such systems tend to be more flexible and dynamic in character as they are continuously shaped by the various stakeholders. Latour believes that such systems cannot be studied by using a hierarchical approach, and proposes network analysis. AT theory implicitly uses such an approach and has been used in conjunction with Latour's own actor-network theory (Latour 1987, 1988) to represent the human and non-human actors of the system (Engeström, 1996b).

In contrast to systemic approaches which presume a minimum of stable structure, AT focuses on the continuous development of the activity system and employs the notion of 'contradiction' (Engeström 1987,1991) as a tool to describe it.

Although Activity Theory is a promising approach for the analysis of systems close to pole B, it doesn't provide a coherent framework to deal with cooperative work. Activity Theory was developed as a psychological approach, and as such it almost exclusively deals with individual human beings (Kaptelinin 1996). There has been an attempt to extend AT through the concept of "collective subject", but as it has been pointed out (Davydov 1999), the relation of individual and collective activity or subject in AT is vague. To overcome such limitations, Engeström uses notions from the actor – network theory, which does provide a basis to analyze collaboration between stakeholders (individuals or groups) but cannot cope with fine grain analysis of cooperative work.

The application of AT has been limited in domains such as medical practice (Engeström 1993, 1995), juridical praxis (Engeström 1996a), programming (Holland et al. 1996), software design (Gould et al, 2000; Turner & Turner, 2001) etc. In such domains it has offered rich accounts mainly of ethnographic character. While such accounts provide useful insights for design, there are few applications of AT on the design of tools supporting human work. This is mainly due to its inherent difficulty i) to provide concrete recommendations at the level needed by designers, ii) to represent the causality that is manifested by the technical components of socio-technical systems.

DISCUSSION AND CONCLUSIONS

This partial account of different cognitive engineering approaches was elaborated in order to contribute to the debate in the cognitive ergonomics community as for the adequacy of different theoretical concepts for the analysis of socio-technical systems in view of designing artifacts to facilitate human work.

To tackle the issue, a taxonomy of socio-technical systems was sketched between two extremes (poles). Close to pole A are systems whose behavior is constrained by a well defined and more or less stable transformation function that reflects the systems' objectives. Close to pole B are systems that their behavior is either governed by multiple objectives and/or their transformation function is vague and subject to important alterations within the lifespan of the system.

The presented approaches share some general characteristics as they all use as unit of analysis some sort of system that encompasses a specific context and rely on extensive field investigations. They reject a normative representation of the work domain and adopt a dynamic view that is framed by some kind of structure or theoretical construct. There are however important differences between them, which reflect to some degree the different theoretical backgrounds but also the positioning of their typical application domains along the above mentioned taxonomy.

The pre-assumption of causality manifested by the abstraction hierarchy in Rasmussen's Cognitive Systems Engineering approach, seems to be influenced by the domain in which it was originally destined to study (e.g. industrial process plants) and by the scope and expected outcomes by its implementation (i.e. design recommendations for supervisory control systems). In the same manner, the Distributed Cognition approach although based on a very different theoretical background (Ethno-methodology), adopts a causal model for its micro-analyses of cognitive systems which is well manifested by its recognition of symmetry between human agents and artifacts.

The two previous approaches presuppose or recognize that the systems they examine can be more or less causally modeled. Therefore, we support that they are appropriate for the analysis of systems attracted by pole A, because in such systems a stable transformation function can be identified.

Approaches based on Activity Theory even though they share some common theoretical ground with Hutchins' Distributed Cognition, clearly differentiate between human actions and the mediating role of artifacts, as they place more emphasis on human intentionality with its consideration of issues such as consciousness and motivation. Thus these approaches reject the predominance of the transformation function as a frame for the systems they study. This is partly due to the nature of the domains studied (e.g. medical practice).

Since the aim is intervention, the effectiveness of these approaches for design must be considered. Cognitive Systems Engineering offers a rigorous methodology with multiple levels of constraint identification in which alternative technological or organizational solutions may be represented. Specifically for information systems it suggests a design grid that maps different constraint types to different phases of design. The strong point of this approach when applied for design is that its hierarchical model of the work domain is compatible with representations of technological artifacts. Thus in principle, future artifacts can be adequately integrated in the models for prediction purposes. On the other hand, as discussed above, Cognitive Systems Engineering is far too rigid to effectively represent systems demonstrating behavior close to pole B. As a result the approach loses its advantages when either the domain or the anticipated intervention has strong social – organizational character.

Distributed Cognition demonstrates a shallower structure than Cognitive Systems Engineering and uses a more flexible and opportunistic unit of analysis. As such, it does not provide any concrete design methodology and at least in its original form Distributed Cognition had more descriptive power than predictive one. Nevertheless it does infer a number of principles for design such as adaptation, cognitive ecology and attention to external representational media (Hutchins, 1995). Additionally its symmetrical consideration of people and artifacts is well adapted for systems suitable to be functionally represented as a large number of local interactions and where internal representations do not play a dominant role. Distributed Cognition may effectively be used in such systems for the comparison of alternative solutions regarding the transfer/transformation of information across agents.

Activity Theory has been widely criticized for its limited usefulness for the design of artifacts. According to Nardi, AT isn't a strong predictive theory but a descriptive tool (Nardi, 1996; Waern, 1998). While AT is by its very nature weak in this respect, there have been voices suggesting that since an accurate prediction of the future activity is impossible, the strength of AT in its acceptance of this impossibility. Thus, for systems close to pole B, the unpredictable future system behavior is dealt with during the analysis and not as a side effect (Bødker, 1991).

Finally it is important to emphasize that the distinction of systems in the proposed A-B pole taxonomy is not claimed to be of an ontological character. Different domains do tend to lie towards one or the other pole but the specific subsystem (i.e. closure) under consideration may demonstrate characteristics that are diverging considerably from its "parent domain". Thus the scope of the analysis and the delimiting of the work domain of interest may play an equally important role in choosing among different approaches. Following a pragmatic point of view the cognitive ergonomist may see the above alternative approaches as complementary tools that may even be used in conjunction to highlight different aspects in a particular intervention.

REFERENCES

Bødker, S. (1991). Activity theory as a challenge to systems design. In Nissen, HE., Klein, H., & R. Hirschheim: *Information Systems Research: Contemporary Approaches and Emergent Traditions*, North Holland, pp. 551-564.

Checkland, P (1981). *Systems thinking, systems practice*. Chichester: John Wiley & Sons.

Davydov, V.V. (1999). The content and unsolved problems of activity theory. In: Engestrom, Y., Miettinen, R., & Punamaki, R.L. (Eds.). *Perspectives on activity theory*, Cambridge University Press, Cambridge.

Engeström, Y. (1987). *Learning by Expanding: An Activity-Theoritical Approach to Developmental Research*. Helsinki: Orienta-Konsultit Oy.

Engeström, Y. (1993). Developmental studies of work as a test bench of activity theory: the case of primary care medical practice. In: S. Chaiklin and J. Lave, (Eds.),*Understanding Practice: Perspectives on Activity and Context.* Cambridge: Cambridge University Press, 64-103.

Engeström, Y. (1995). Objects, contradictions and collaboration in medical cognition: an activity-theoretical perspective. *Artificial Intelligence in Medicine*, 7, 395-412

Engeström, Y. (1996)a. The Tensions of Judging. In: Engeström, Y. & Middleton, D. (Eds). *Cognition and Communication at Work*. Cambridge: MIT Press

Engeström, Y. (1996)b. Mundane Tool or Object of Affection? The Rise and Fall of the Postal Buddy. In Nardi, B.A. (Ed), *Context and Consciousness*. Cambridge: MIT Press.

Gould, E., Verenikina, I and Hasan H. (2000). Activity Theory as a Basis for the Design of Web Based System of Inquiry for World War 1 Data. In: *Proceedings of the 23rd IRIS Conference*. Lingatan, Sweden, August, 2000, pp. 761-770.

Hollan, J.D., Hutchins, E. & Kirsch, D. (2001). Distributed cognition: a new foundation for human-computer interaction. *ACM Trans. Computer-Human Interaction, 7,* 2, 174-196

Holland, D. & Reeves J.R. (1996). Activity Theory and the View from Somewhere: Team Perspective on the Intellectual Work of Programming. In Nardi, B.A. (Ed) (1996). *Context and Consciousness*. Cambridge: MIT Press.

Hollnagel, E. (2001). Extended cognition and the future of ergonomics. *Theoretical Issues in Ergonomics Science,* 2(3), pp. 309-315.

Hutchins E. & Klausen, T. (1996). Distributed cognition in an airline cockpit. In: Y. Engestrφm and D. Middeleton (Eds.). *Cognition and Communication at Work*, Cambridge University Press, 15-34.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge: MIT press.

Kirsh, D. (2001). The context of work. *Human-Computer Interaction,* 16.

Kuutti, K. (1996). Activity Theory as a Potential Framework for Human-Computer Interaction Research. In: Nardi, B.A. (Ed) (1996). *Context and Consciousness*. Cambridge: MIT Press

Latour, B. (1987). *Science in Action*. Cambridge, MA: Harvard University Press.

Latour, B. (1988). *The Pasteurization of France*. Cambridge, MA: Harvard University Press

Lemke, J.L. (2000). Material Sign Processes and Ecosocial Organization. In: P.B. Andersen, C. Emmeche, and N.O. Finnemann-Nielsen, (Eds.), *Downward Causation: Self-organization in Biology, Psychology, and Society*. Aarhus University Press (Denmark). pp. 181-213. 2000.

Leontiev, A. (1974). The problem of activity in psychology. *Soviet Psychology*, 13 (2), 4-33.

Nardi, B.A. (1996). *Context and Consciousness*. Cambridge, MA: MIT Press

Perrow, C. (1984). *Normal accidents: Leaving with high-risk technologies*. Princeton: Princeton University Press. (2nd edition)

Petersen, J. & Nielsen, M. (2001). Analyzing Maritime Work Domains. In: Onken, R. (Ed.) Proc. of 8th *Cognitive Science Approaches to Process Control*. Sept. 24-26, 2001, Neubiberg, Germany.

Rasmussen, J., Pejtersen, A.M. & Goodstein, L.P. (1994). *Cognitive systems engineering*. NY: John Wiley & Sons.

Sharp, T. D., Helmicki, A. J. (1996). Applying the Abstraction Hierarchy to Intensive Care Medicine. In: *Proceedings of the Symposium on Human Interaction with Complex Systems-96*, August, 1996 .

Simon, H.A. (1976). *Administrative behavior*. NY: The Free press, 3rd edition.

Turner, P. & Turner,S. (2001). Describing Team Work With Activity Theory. *Cognition, Technology & Work* (2001) 3:127–139. Springer-Verlag

Waern, Y., (1998). Cognitive Ergonomics of Collaborative Work. Presented in the 9th European Conference on Cognitive Ergonomics.

Yoon, B.-d. & Garcia, O. N. (1996). Hierarchical Problem-Solving in Software Debugging. In *Proceedings of the 3rd Symposium on Human Interaction with Complex Systems-96* (HICS '96), Dayton, OH. August, 1996 .