

# From Design Experiments to Formative Interventions

Yrjö Engeström, University Of Helsinki, Center for Activity Theory and Developmental Work Research  
P.O.Box 26, 00014 University of Helsinki, Finland, yrjo.engestrom@helsinki.fi

Anyone who attempts to skip this problem, to jump over methodology in order to build some special psychological science right away, will inevitably jump over his horse while trying to sit on it.  
(Vygotsky, 1997a, p. 329)

The reputation of educational research is tarnished less by the lack of replicable results than by the lack of any deeper theory that would explain why the thousands of experiments that make up the literature of the field appear to have yielded so little.  
(Olson, 2004, p. 25)

**Abstract:** The discussion of design experiments has largely ignored the Vygotskian tradition of formative interventions based on the principle of double stimulation. This tradition offers a radical approach to learning research which focuses on the agency of the learners. The principle of double stimulation is used and developed further in the intervention methodology called Change Laboratory, created in the Center for Activity Theory and Developmental Work Research at University of Helsinki. The paper analyzes the Change Laboratory methodology and its potential for generating expansive learning, using data from an intervention conducted in 2006 in the surgical unit of a university hospital in Finland. The analysis demonstrates how the agency of the learning collective developed hand-in-hand with the construction and implementation of a new organization of work by the collective. Such expansive learning goes beyond knowledge construction, resulting in materially anchored new practices.

## Introduction

Human learning takes place within and between complex, continuously changing activity systems. Learning needs themselves are increasingly opaque. It is not at all clear just what needs to be learned to cope with the demands of complex activities and global networks in constant turmoil. Humans – practitioners, teachers, students – are intentional and interactive beings who keep interpreting and reinterpreting the challenges and tasks they face in their own, multiple, changing, and often unpredictable ways. They do not neatly obey the laws of linear causality. The practical usefulness and ecological validity of research on learning based on classical well controlled experiments are more questionable than ever.

In the past few years, the United States educational authorities have aggressively launched legislation and national guidelines that define the ‘gold standard’ of educational research. The ‘gold standard’ emphasizes the use of randomized controlled trials, the selection of valid control groups, and ‘scalability’ implying large statistical samples and multiple research sites.

The ‘gold standard’ correctly sees educational research as interventionist research. The randomized control trials are meant to assess the effectiveness of educational interventions. The model of intervention research is taken from fields such as medicine and agriculture. As one observer put it:

“For instance, if I want to test the effectiveness of weed control measures, I randomly assign different plots of crops to the experimental or control conditions. Then, they all get treated the same otherwise as far as weather, fertilizer, hours of day light and other pests. The crops are monitored and observations are made throughout the growing season and a person might be able to see the result visually if the results are remarkable enough. But the telling evidence is in the yield, when the crops are harvested. If there is a significant difference in yield in all the experimental plots as opposed to the control plots, then we might attribute it towards the independent variable, which in this case is weed control.”

(<http://specialized.wordpress.com/2006/02/10/educational-researchthe-gold-standard/>)

The ‘gold standard’ thinking in educational research starts from the assumption that researchers know what they want to implement, how they want to change the educational practice. In other words, the intervention and its desired outcomes are well defined in advance. The task of research is to check whether or not the desired outcomes are actually achieved.

Educational researchers are in a bind. On the one hand, many of them recognize the limits of randomized control trials and seek ways to conduct and legitimize more practice-based and creative and theoretically ambitious research. On the other hand, there are strong administrative, financial, political and

'scientific' pressures to stick to the proven assumptions and methodological rules of positivist science. It is no wonder that many attempts at methodological innovation turn out to be weak compromises.

In this paper, I will propose a radical methodological approach that I call formative interventions. The approach is inspired by the methodological ideas and experiences of L. S. Vygotsky and his colleagues and students. My version of this legacy has been worked out over a period of some 20 years of research in the Finnish school of developmental work research (Engeström, 2005, Engeström, Lompscher & Rückriem, 2005).

I will begin by examining the recent interest in so called design experiments or design research, which I find to be still captive of the linear view of interventions typical to the 'gold standard'. After that, I will discuss the contribution of some sociological analyses of interventions, pointing toward the crucial importance of agency in developing a viable interventionist methodology. I will present Vygotsky's central ideas of the interventionist methodology called 'double stimulation', and conclude by pointing out three key differences between linear and formative interventions. This leads me to present 'the argumentative grammar' of formative interventions with the help of four foundational epistemic threads that inform the methodology: 1) units of analysis, 2) causality and agency, 3) change and development, and 4) generalization.

I will illustrate the methodology of formative interventions by analyzing data from an intervention study my research group recently carried out in Finland with the help of the Change Laboratory toolkit. This study concerns the formation of a new mode of working in the central surgery unit of the Oulu University Hospital. At the end of the paper, I will discuss the possibilities of using formative interventions in school education.

## Design experiments – Design research

In this section, I will take up problematic aspects in recent literature on design experiments, or design research. I will not go back to the early formulations of this research approach (Brown, 1992, Collins, 1992), as their main arguments are widely known.

In literature on design experiments or design research the focus of attention is moved from isolated individual learners to entire learning environments or learning ecologies.

"Design experiments ideally result in greater understanding of a *learning ecology* - a complex, interacting system involving multiple elements of different types and levels - by designing its elements and by anticipating how these elements function together to support learning." (Cobb & al, 2003, p. 9)

Many proponents of design research use the notion of dynamical learning environments (DLEs) to describe their unit of analysis. For example Barab and Kirshner (2001, p. 8) define dynamical learning environments "as continually evolving as the system components (students, teachers, class assignments, resources) that constitute DLEs reciprocally interact in ways that both stabilize and destabilize the system."

These notions of complex units of analysis suffer from vagueness. Talking in general terms about systems, dynamics and components is not enough. What exactly might be a useful model of the anatomy of such a dynamic learning environment? If design researchers do not specify and model the crucial components and relations of their proposed units of analysis, a decisive connection between theory and methodology is severed.

In discourse on 'design experiments', it seems to be tacitly assumed that researchers make the grand design, teachers implement it (and contribute to its modification), and students learn better as a result. Scholars do not usually ask: Who does the design and why? This linear view is associated with notions of perfection, completeness and finality. This is exemplified by the use of the absurd notion of capturing 'all variables'.

"In design experiments there are many dependent variables that matter, though the researchers may not pay attention to them *all*. (...) the goal is to identify *all* the variables, or characteristics of the situation, that affect any dependent variables of interest." (Collins, Joseph & Bielaczyc, 2004, p. 20; emphasis added by Y.E.)

As a byproduct, Collins, Joseph and Bielaczyc endorse a variable-oriented approach to research without questioning the extremely problematic notion of causality behind it (Maxwell, 2004). In the account of Collins, Joseph and Bielaczyc (2004, p. 33), the methodology of design research is basically a linear progression of six steps, starting from 'implementing a design' and ending with 'reporting on design research'. As the process begins with implementation, the making of the design in the first place is not even included in the methodology. Thus, there is no need to problematize the issue of who makes the design and guided by what theory or principles. In a similar vein, Cobb and his co-authors (2003) seem to take it for granted that it is the researchers who determine the 'end points' for the design experiment.

"In addition to clarifying the theoretical intent of the experiment, the research team must also

specify the significant disciplinary ideas and forms of reasoning that constitute the prospective goals or endpoints for student learning.” (Cobb & al., 2003, p. 11)

The stepwise linear notion of design research is also explicated by Bannan-Ritland (2003, p. 22). Cyclic iterations serving the refinement of the design complement but do not challenge the basically linear image.

This linear view ignores what we know of interventions as contested terrains, full of resistance, reinterpretation and surprises from the actors below. Cobb and co-authors do mention that design experiments conceived by researchers create discontinuity – but that does not seem to require any further reflection.

”The intent is to investigate the possibilities for educational improvement by bringing about new forms of learning in order to study them. Consequently, there is frequently a significant discontinuity between typical forms of education (these could be studied naturalistically) and those that are the focus of a design experiment.” (Cobb & al., 2003, p. 10)

The main difference between ‘gold standard’ interventions and design experiments seems to be that the former expects the design of the intervention to be complete at the outset while the latter, recognizing the complexity of educational settings, expects the design to proceed through multiple iterations of ‘refinement’. But even design experiments aim at closure and control. The emphasis on completeness, finality and closure is condensed in the idea of design experiments as ‘refinement’. The implication is that the researchers have somehow come up with a pretty good model which needs to be perfected in the field.

”Design experiments were developed as a way to carry out formative research to test and refine educational designs based on theoretical principles derived from prior research. This approach of progressive refinement in design involves putting a first version of a design into the world to see how it works. Then, the design is constantly revised based on experience, until *all the bugs* are worked out.” (Collins, Joseph & Bielaczyc, 2004, p. 18; emphasis added by Y.E.)

”Design research should always have the dual goals of refining both theory and practice.” (Collins, Joseph & Bielaczyc, 2004, p. 19)

Collins, Joseph and Bielaczyc (2004, p. 18-19) compare educational design research to the design of cars and other consumer products, using *Consumer Reports* as their explicit model for evaluation. They don’t seem to notice any significant difference between finished mass products and such open-ended, continuously co-configured products as educational innovations (for co-configuration, see Victor & Boynton, 1998, Engeström, 2004b). A strange obsession with ‘completeness’ runs like a red thread through their argument.

”Thus, in the jigsaw, all pieces of the puzzle come together to form *a complete understanding*.” (Collins, Joseph & Bielaczyc, 2004, p. 23; emphasis added by Y.E.)

What this overlooks is that “one can never get it right, and that innovation may best be seen as a continuous process, with particular product embodiments simply being arbitrary points along the way” (von Hippel & Tyre, 1995, p. 12).

To sum up, recent literature proposing design experiments or design research as a methodology for educational research suffer from serious built-in weaknesses. While there are many different versions of design research, it seems fair to conclude that the following weaknesses are quite pervasive. First, the unit of analysis is left vague. Secondly, the process of design research is depicted in a linear fashion, starting with researchers determining the principles and goals and leading to completion or perfection. This view ignores the agency of practitioners, students and users, as well as the crucial difference between finished mass products and open-ended social innovations. Finally, in much of the literature on design experiments, a variable-oriented approach to research is tacitly endorsed, without questioning the underlying problematic notion of causality (I will return to this issue shortly).

## **Lessons from sociological intervention research**

Sociological intervention studies differ from educational ones in that there are usually no safe institutional walls to protect the intervention from the vagaries of the outside world. Perhaps this is why the linear view I observed in design research is much less easily adopted in sociology. A good case in point is the work of Norman Long (2001).

“Intervention is an on-going transformational process that is constantly re-shaped by its own internal organisational and political dynamic and by the specific conditions it encounters or itself creates, including the responses and strategies of local and regional groups who may struggle to define and defend their own social spaces, cultural boundaries and positions within the wider power field.” (Long, 2001, p. 27)

Long uses words like struggle, strategy, power and position – words that are conspicuously absent in recent literature on design experiments.

“Crucial to understanding processes of intervention is the need to identify and come to grips with the strategies that local actors devise for dealing with their new intervenors so that they might appropriate, manipulate, subvert or dismember particular interventions.” (Long, 2001, p. 233)

In other words, resistance and subversion are not accidental disturbances that need to be eliminated. They are essential core ingredients of interventions, and they need to have a prominent place in a viable intervention methodology. Melucci (1996) extends this point into a threefold methodological guideline for intervention research.

“What we must recognize is that actors themselves can make sense out of what they are doing, autonomously of any evangelical or manipulative interventions of the researcher. (...) Secondly, we need to recognize that the researcher-actor relation is itself an object of observation, that it is itself part of the field of action, and thus subject to explicit negotiation and to a contract stipulated between the parties. (...) Lastly, we must recognize that every research practice which involves intervention in the field of action creates an artificial situation which must be explicitly acknowledged. (...) a capability of metacommunication on the relationship between the observer and the observed must therefore be incorporated into the research framework.” (Melucci, 1996, p. 388-389)

To other words, interventions in human beings’ activities are met with actors with identities and agency, not with anonymous mechanical responses. If agency is not a central concern in the methodology, there is something seriously wrong with it.

In educational research, one of the few scholars who have taken this seriously is David Olson.

“Research in the human sciences, it may be argued, is less designed to dictate what one does than to provide information that agents, both teachers and students, can use in making informed decisions about what to do in the multiple and varied contexts in which they work.” (Olson, 2004, p. 25).

## **Vygotsky’s method of double stimulation**

In his quest for a new psychology based on cultural mediation of higher mental functions, Vygotsky was very conscious of the need to build a methodology that would correspond to the character of the theory.

“This methodology [study of reactive responses based on the S-R formula], which easily establishes the response movements of the subject, becomes completely impotent, however, when the basic problem is the study of those means and devices that the subject used to organize his behavior in concrete forms most adequate for each given task. In directing our attention to the study of specifically these (external and internal) means of behavior, we must conduct a radical review of the methodology of the psychological experiment itself.” (Vygotsky, 1999, p. 59)

The methodology Vygotsky, Leont’ev and Luria developed has been characterized by different names. Vygotsky (e.g., 1997b, p. 68, 1997c, p. 85-89, 1999, p. 57-59) used at least the names ‘experimental-genetic method’, ‘instrumental method’, ‘historical-genetic method’, and ‘method of double stimulation’, somewhat interchangeably. In this paper, I will use the ‘method of double stimulation’.

As van der Veer and Valsiner (1991, p. 169) put it, in double stimulation experiments, “the subject is put in a structured situation where a problem exists (...) and the subject is provided with active guidance towards the construction of a new means to the end of a solution to the problem.” Vygotsky himself described the methodology as follows.

“The task facing the child in the experimental context is, as a rule, beyond his present capabilities and cannot be solved by existing skills. In such cases a neutral object is placed near the child, and frequently we are able to observe how the neutral stimulus is drawn into the situation and takes on the function of a sign. Thus, *the child actively incorporates these neutral objects into the task of problem solving*. We might say that when difficulties arise, neutral stimuli take on the function of a sign and from that point on the operation’s structure assumes an essentially different character.” (Vygotsky, 1978, p. 74; italics added)

“By using this approach, we do not limit ourselves to the usual method of offering the subject simple stimuli to which we expect a direct response. Rather, we simultaneously offer a *second series of stimuli* that have a special function. In this way, we are able to study the *process of accomplishing a task by the aid of specific auxiliary means*; thus we are also able to discover the inner structure and development of higher psychological processes.

The method of double stimulation elicits manifestations of the crucial processes in the behavior of people of all ages. Tying a knot as a reminder, in both children and adults, is but one example of a pervasive regulatory principle of human behavior, that of *signification*, wherein people create temporary links and give significance to previously neutral stimuli in the context of their problem-solving efforts. We regard our method as important because it helps to *objectify* inner psychological processes...” (Vygotsky, 1978, p. 74-75)

It is important to note that the second stimuli, the mediating means, were not necessarily given to the subjects in any ready-made form.

“In experimental studies, we do not necessarily have to present to the subject a prepared external means with which we might solve the proposed problem. The main design of our experiment will not suffer in any way if instead of giving the child prepared external means, we will wait while he spontaneously applies the auxiliary device and involves some auxiliary system of symbols in the operation. (...) In not giving the child a ready symbol, we could trace the way all the essential mechanisms of the complex symbolic activity of the child develop during the spontaneous expanding of the devices he used.” (Vygotsky, 1999, p. 60)

Van der Veer and Valsiner (1991, p. 399) point out the fundamental challenge this methodology poses to the experimenter who wants to control the experimental situation.

“The notion of ‘experimental method’ is set up by Vygotsky in a methodological framework where the traditional norm of the experimenter’s maximum control over what happens in the experiment is retained as a special case, rather than the modal one. The human subject always ‘imports’ into an experimental setting a set of ‘stimulus-means’ (psychological instruments) in the form of signs that the experimenter cannot control externally in any rigid way. Hence the experimental setting becomes a context of investigation where the experimenter can manipulate its structure in order to trigger (but not ‘produce’) the subject’s *construction* of new psychological phenomena.”

In other words, the subject’s agency steps into the picture. To fully appreciate the radical potential of the methodology of double stimulation, we need to reconstruct Vygotsky’s more general conception of intentionality and agency. Vygotsky described this artifact-mediated nature of intentional action as follows.

“The person, using the power of things or stimuli, controls his own behavior through them, grouping them, putting them together, sorting them. In other words, the great uniqueness of the will consists of man having no power over his own behavior other than the power that things have over his behavior. But man subjects to himself the power of things over behavior, makes them serve his own purposes and controls that power as he wants. He changes the environment with the external activity and in this way affects his own behavior, subjecting it to his own authority.” (Vygotsky, 1997b, p. 212)

Vygotsky (1997b, p. 213) pointed out that voluntary action has two phases or ‘two apparatus’. The first one is the design phase in which the mediating artifact or “the closure part of the voluntary process” is, often painstakingly, constructed. The second one is the execution phase or “actuating apparatus” which typically looks quite easy and almost automatic, much like a conditioned reflex.

Classic examples of culturally mediated intentionality include devices we construct and use to wake up early in the morning. Vygotsky's examples of voluntary action are mostly focused on individual actors. This must not be interpreted as neglect of collective intentionality. According to Vygotsky's famous principle, higher psychological functions appear twice, first interpsychologically, in collaborative action, and later intrapsychologically, internalized by the individual.

"V. K. Arsen'ev, a well-known researcher of the Ussuriysk region, tells how in an Udeg village in which he stopped during the journey, the local inhabitants asked him, on his return to Vladivostok, to tell the Russian authorities that the merchant Li Tanku was oppressing them. The next day, the inhabitants came out to accompany the traveler to the outskirts. A gray-haired old man came from the crowd, says Arsen'ev, and gave him the claw of a lynx and told him to put it in his pocket so that he would not forget their petition about Li Tanku. The man himself introduced an artificial stimulus into the situation, actively affecting the processes of remembering. Affecting the memory of another person, we note in passing, is essentially the same as affecting one's own memory." (Vygotsky, 1997b, p. 50-51)

Vygotsky's colleague A. N. Leont'ev (1932) focused on the social origins of intentional action. He pointed out that signals given by foremen, the rhythmic sounds of a drum, and working songs gave collective work the necessary direction and continuance. The interpsychological origins of voluntary action – and collective intentionality – would thus be found in rudimentary uses of shared external signals, prompts, as well as in reminders, plans, maps, etc.

We see the radical potential of double stimulation and mediated intentionality every day in educational practice. Cheating in school is an enlightening example. What does a student do when she constructs a cheating slip while preparing for an exam?

The exam questions and the texts one must master are the 'first stimuli', or the object, for the student. The cheating device, for example a paper slip, is the 'second stimulus', or the mediating tool. The cheating slip is typically a small piece of paper that can be hidden away from the teacher's eyes and on which one writes what one considers the most essential information about a topic one expects to be included in the exam questions. Since the slip is small, there cannot be too much text. To create a good cheating slip, the student must carefully select the most relevant and useful aspects of the topic and represent them in an economic and accessible way on the slip. Thus, the construction of a cheating slip is truly what Vygotsky described as creating an external auxiliary means for mastering an object. The construction, contents and use of the cheating slip bring into light and objectify the inner psychological process of preparing for the test. If we get access to the construction, contents and use of cheating slips we learn much more about students' learning than merely by reading and grading their exam answers. That is why I occasionally ask my students to prepare cheating slips and to cheat in my exam, then at the end of the exam I collect their slips and the actual answers.

Cheating is an important form of student agency. By creating and using a cheating slip, the student controls his or her own behavior with the help of a tool he or she made. The hard part is the construction of a good cheating slip – the design phase or the 'closure part' of the agentic action. When asked, students often report that the execution part is surprisingly easy. If the slip has been well prepared, it is often enough that the student merely glances at it – the details seem to follow from memory as if a floodgate had been opened. This is the phenomenon of instantaneous recollection or reconstruction of a complex meaningful pattern with the help of a good 'advance organizer' (Ausubel, 2000), 'orientation basis' (Haenen, 1995, Talyzina, 1981), or 'germ cell model' (Davydov, 1990). In other words, learning to cheat well is extremely valuable.

At the same time, cheating is contestation of the given activity of school-going. By constructing and using a cheating slip, the student takes a risk but also creates a new mediating tool for the mastery of the entire testing situation, which is really the core of traditional schooling. This goes far beyond merely quantitatively enlarging or 'amplifying' one's memory. Good cheating is a way to beat the system, to be more clever than the given activity. Long ago John Holt (1964) gave a vivid picture of the beginnings of this type of agency when he described how elementary school kids learn to calculate the risk: When the teacher asks a question to which you don't know the answer, it is reasonably safe to raise your hand if most of the other kids also raise their hand. You look good and the probability of getting caught is low. Agency is by definition testing and going beyond the limits of what is required and allowed. Students are themselves making double-stimulation experiments in these situations.

Vygotsky's methodological principle of double stimulation leads to a concept of formative interventions which are radically different from the linear interventions advocated both by the 'gold standard' and by the literature on design experiments. 1) The crucial differences may be condensed in three points.

- 1) *Starting point:* In linear interventions, the contents and goals of the intervention are known ahead of time by the researchers. In formative interventions, the subjects (whether children or adult

practitioners, or both) face a problematic and contradictory object which they analyze and expand by constructing a novel concept, the contents of which are not known ahead of time to the researchers.

- 2) *Process*: In linear interventions, the subjects (typically students in school) are expected to receive the intervention without argument. Difficulties of reception are interpreted as weaknesses in the design that are to be corrected. In formative interventions, the contents and course of the intervention are subject to negotiation and the shape of the intervention is eventually up to the subjects. Double stimulation as the core mechanism implies that the subjects gain agency and take charge of the process.
- 3) *Outcome*: In linear interventions, the aim is to control all the variables and to achieve a standardized solution module (typically a new learning environment) that will reliably generate the same desired outcomes when transferred and implemented in new settings. In formative interventions, the aim is to generate intermediate concepts and solutions that may be used in other settings as tools in the design on locally appropriate new solutions.

Intervention may be defined simply as “purposeful action by a human agent to create change” (Midgley, 2000, p. 113). This definition makes it clear that the researcher does not have a monopoly over interventions. Organized activity systems such as schools and workplaces are bombarded by interventions from all kinds of outside agents (consultants, administrators, customers, competitors, partners, politicians, journalists). And inside the activity system, practitioners and managers incessantly make their own interventions. We as researchers should not expect nicely linear results from our efforts.

## **Toward an argumentative grammar**

In his discussion of design research in education, Kelly (2004) took up the foundational issue of an ‘argumentative grammar’ that is needed to formulate a viable methodology.

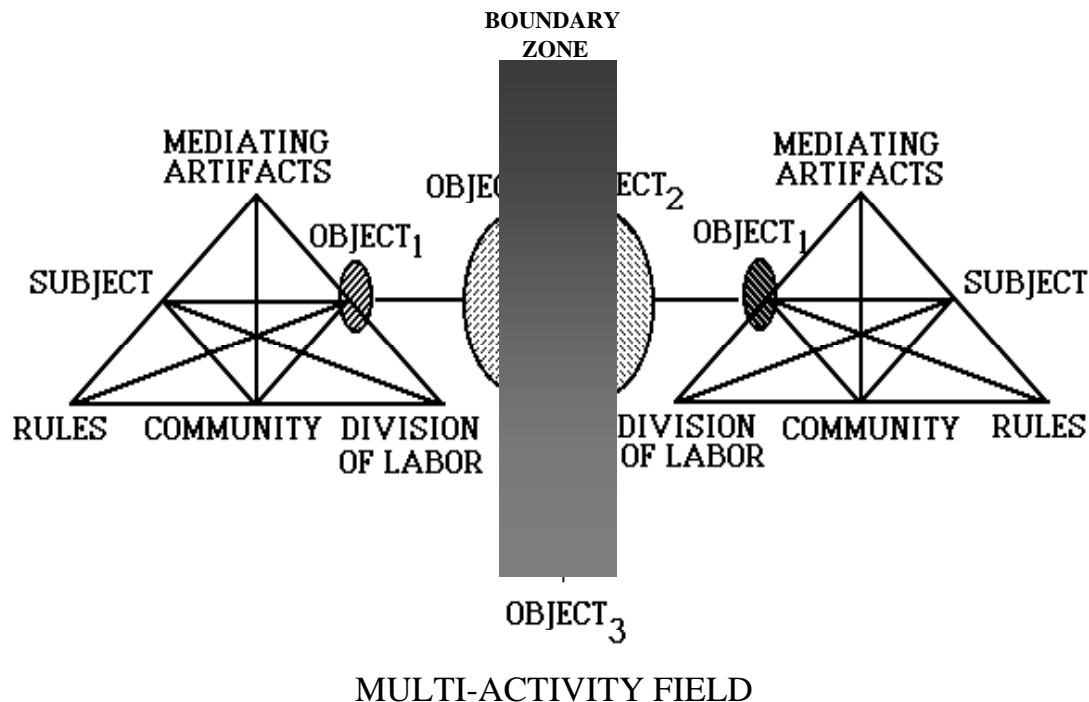
“An argumentative grammar is the logic that guides the use of a method and that supports reasoning about its data. It supplies the *logos* (reason, rationale) in the methodology (method + logos) and is the basis for the *warrant* for the claims that arise.” (Kelly, 2004, p. 118)

I see the argumentative grammar as a set of basic epistemic ideas or threads that run through and connect theory, methodology, and empirical research in any serious research approach. Such epistemic ideas are at the same time both substantive assumptions about the nature of the objects of research and heuristic meta-level tools for the practical conduct of intervention, data collection and analysis.

“What, therefore, is the logos of design studies in education? What is the grammar that cuts across the series of studies as they occur in different fields? Where is the ‘separable’ structure that justifies collecting certain data and not other data and under what conditions? What guides the reasoning with these data to make a plausible argument?” (Kelly, 2004, p. 119)

I will suggest four epistemic threads as a guiding framework for presenting and examining the methodology of formative interventions: (1) unit of analysis, (2) causality and agency, (3) change and development, and (4) generalization. These four are certainly not the only possible or exhaustive set of basic epistemic questions.

1) *Units of analysis*. We may discern three generations in the evolution of the prime unit of analysis within cultural-historical activity theory (Engeström, 2001). The first generation, based on Vygotsky’s work, centered around mediated action as unit of analysis (see Zinchenko, 1985). The second generation, based on Leont’ev’s (1978, 1981) work, took the collective activity system as its molar unit of analysis. At present, many activity theorists in various parts of the world are focusing on interactions among two or more activity systems, which requires a third generation unit of analysis where minimally two activity systems have a partially shared object (Figure 1).



**Figure 1.** Third generation activity theory: two activity systems with a partially shared objects as minimal unit of analysis.

The three successive units are not mutually exclusive. The mediated action of the 1st generation may be recognized as the tip of the iceberg in the 2nd and 3rd generation models. More importantly, a mature view of the unit of analysis may actually require multiple levels, as was already forcefully pointed out by Leont'ev (1978). In other words, while one may want to focus, say, on an activity system as the prime unit of a study, this unit only gains explanatory power if one makes visible transitions and boundaries between the activity system and the actions it generates on the one hand, and between the activity system and the field of interconnected activity systems in which it is located on the other hand.

The unit of analysis is usually a conceptual idea strictly for the researchers. In formative interventions, we turn the unit of analysis into an external auxiliary means, a mediating conceptual tool, for both the participating subjects and the researchers. The triangular models of activity are typically presented and explained to the participants at an early phase of the intervention. Thereafter, they are repeatedly used to analyze and redesign various aspects of the participants' activity. In other words, the unit of analysis itself is offered as a potential 'cheating slip' to the participants.

As is evident in Figure 1, activity theory sees the object as the crucial factor that gives durable direction, purpose and identity to an activity. As activity systems become fragmented by their inner contradictions, the object of activity tends to get blurred or lost. The methodology of formative interventions is oriented at the rediscovery and expansive reforging of the object of activity.

An important aspect of activity-theoretical units of analysis is the constant movement between the long historical time perspective of object-oriented activity systems and the relatively short time perspective of goal-oriented actions (Engeström, 2000). Development, understood as qualitative transformations, is driven by an expansive reconceptualization of the object and motive of the entire *activity*. But such transformations are both initiated and implemented in daily work *actions*, in deviations from the prescribed course of actions and in mundane innovations. The crucial issue is movement between these two levels.

Talk at the level of entire activity systems is typically saturated by envisioning and imagining, characterized by hypothetical and tentative formulations ('what if?'), as well as efforts at conceptualization and modeling of principles or general rules. Discussion at this level bears references to issues that go beyond the here-and-now both in time and in space. Talk at the level of actions is typically focused on a particular object or problem demanding decisions or solutions. We might say that activity-level discourse generates *visions*, while action-level discourse generates consequential *decisions*. If the two are kept separate, visions tend to remain daydreams and decisions tend to be reactive rather than proactive (Engeström, 2004a, Engeström, Engeström & Kerosuo, 2003).

The first generation unit introduces the concepts of subject, object and mediating artifact as constituents of action. The second generation model proposes an anatomy of collective activity, adding the




concepts of community, rules, and division of labor, as well as the sub-processes of production, distribution, exchange, and consumption, and finally also sense and meaning. The third generation model goes beyond the interplay between action and activity. It introduces the concepts of multi-activity field or terrain and boundary zone, aimed at capturing crucial features of units in which two or more activity systems are connected by a partially shared object.

2) *Causality and agency*. Maxwell (2004) calls the traditional notion of causality the ‘regularity’ approach. It holds that we cannot directly observe causation, only the regularities in the relationships between events. The regularity approach necessarily entails a variable-oriented view of research. Causation is understood as a systematic relationship between variables rather than a causal process. In contrast to variable-based research, process-oriented research believes that causation can actually be observed and reconstructed as a real sequence of events. It uses historical methods and narrative evidence, as well as close observation and recording of unfolding chains of events.

But how does one observe and reconstruct chains of events among human beings? What kind of interpretive lenses do we need for that? Eskola (1999, p. 111) suggests that the answer lies in three facets: (1) the structure and development of the activity in which the actors are involved and its meaning to the different actors, (2) the laws and rules that actors take into account in this activity, and (3) the logics on the basis of which they do so. Eskola presents the basic explanatory schemes of traditional variable oriented research on the one hand and of ‘realistic research in human action’ on the other hand. Eskola’s realistic paradigm focuses on the fact that humans do not merely react as physical objects; they act based upon their activities, interpretations and logics. For the sake of simplicity, we may this the *interpretive layer* of causality.

But there is more to causality in human contexts. Human beings not only interpret, they also face contradictions between multiple motives embedded in and engendered by their historically evolving communities and objects. This is the layer that makes humans look irrational and unpredictable (Engeström, 1989). This adds another layer to human causality. I call it the *contradictory layer*. What is still missing is the human potential for agency, for intentional collective and individual actions aimed at transforming the activity. Thus, I complete the picture by adding an *agentive layer* (Figure 2).

INTERPRETIVE LAYER	IN ACTIVITY THE ACTOR	TAKES INTO ACCOUNT, ACCORDING TO THIS OR THAT LOGIC, THAT	IF X, THEN Y  LAW, RULE
CONTRADICTIONARY LAYER	AS PARTICIPANT IN COLLECTIVE ACTIVITIES	IS DRIVEN BY CONTRADICTIONARY MOTIVES	SEARCHING RESOLUTION BY OFTEN UNPREDICTABLE ACTIONS
AGENTIVE LAYER	AS POTENTIAL INDIVIDUAL AND COLLECTIVE AGENT	TAKES INTENTIONAL TRANSFORMATIVE ACTIONS	INVENTING AND USING ARTIFACTS TO CONTROL THE ACTION FROM THE OUTSIDE

**Figure 2.** Three layers of causality in human action.

Here I remind the reader of the earlier discussion of Vygotsky’s view of agency as originating in the use of external artifacts to reach a redefinition of a situation.

“In experiments involving meaningless situations, Lewin found that the subject searches for some point of support that is external to him and that he defines his own behavior through this external support. In one set of experiments, for example, the experimenter left the subject and did not return, but observed him from a separate room. Generally, the subject waited for 10-20 minutes. Then, not understanding what he should do, he remained in a state of oscillation, confusion and indecisiveness for some time. Nearly all the adults searched for some external point of support. For example, one subject defined his actions in terms of the striking of the clock. Looking at the clock, he thought: ‘When the hand moves to the vertical position, I will leave.’ The subject transformed the situation in this way, establishing that he would wait until 2:30 and then leave. When the time came, the action occurred automatically. By changing the psychological field, the subject created a new situation for himself in this field. He transformed the meaningless situation into one that had a clear meaning.” (Vygotsky, 1987a, p. 356)

Vygotsky's description of Lewin's experiment captures all the three layers of Figure 2 in a simplified form. Initially, the subject interprets the situation as an experiment in which one must follow the rules of the experimenter. When nothing happens, a contradiction emerges between those expected rules and one's quest for meaning; there is a period of confusion, which could lead to unpredictable and 'irrational' actions. However, by using an external cultural artifact such as the clock, the subject is able to transform the situation and take agentive action. Notice that agentive action in its rudimentary forms may look like non-action, or mere resistance – such as leaving the room in the experiment. It is nonetheless a radically different action from that of, say, passive waiting or 'irrational' making of noise.

3) *Change and development.* Adrian Cussins' (1992, 1993) theory of cognitive trails is an account of embodied cognition where the basic metaphor is that of a person moving in a territory. The key concepts are perspective-dependence (PD) and stabilization. Imagine a person standing somewhere in the middle of a city. The person's ability to find his or her way to any desired location regardless of the person's initial position is called perspective-independence. In such case, the PD ratio is high - close to 1. The PD ratio is close to zero when the person is completely unable to find his or her way to any desired location in the territory.

People learn to move around in a territory by moving around in the territory. In so doing, they make cognitive trails.

"Trails are both person-made and world-made, and what makes persons and worlds. Trails are in the environment, certainly, but they are also *cognitive* objects. A trail isn't just an indentation in a physical surface, but a *marking* of the environment; a signposting for coordinating sensation and movement, an experiential line of force. Hence the marking is both experiential and environmental. " (Cussins, 1992, p. 673-674)

"Each trail occurs over time, and is a manipulation or a trial or an avoidance or capture or simply a movement. It is entirely context-dependent [...]. Yet a trail is not transitory (although a tracking of a trail is): the environmental marking persists and thereby the ability to navigate through the feature-domain is enhanced." (Cussins, 1992, p. 674)

As multiple trails are marked, some trails intersect. Intersections are landmarks. A territory is structured by means of a network of landmarks. Such structuring means increasing the PD ratio.

Along with the PD ratio, there is another dimension that characterizes the development of cognitive trails, namely stabilization. Stabilization may also be characterized as blackboxing.

"Stabilization is a process which takes some phenomenon that is in flux, and draws a line (or builds a box) around the phenomenon, so that the phenomenon can enter cognition (and the world) in a single act of reference [...]." (Cussins, 1992, p. 677)

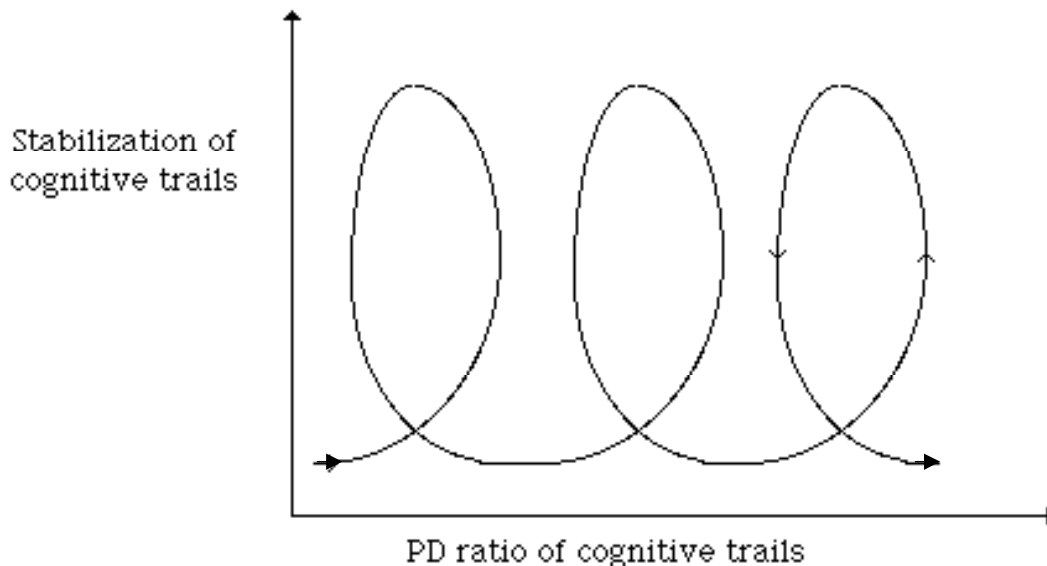
"There comes a time when it is best to stabilize a network of trails so that the space is treated cognitively (functions) as a given unit (an object!), and then build higher-order feature-spaces [...]." (Cussins, 1992, p. 679)

"One familiar and important way in which stabilization is achieved is by drawing a linguistic blackbox around a feature-space: the imposition of linguistic structure on experiential structure. [...] A region of feature-space starts to function as an object as it is dominated by a network of trails and stabilized by a name." (Cussins, 1992, p. 679-680)

The point of maximum generality is where stabilized objects, concepts and explicit propositions emerge. Cussins depicts cognition as "appropriate spiraling" in the two-dimensional terrain described above. He calls this movement "virtuous representational activity" (Figure 3).

"The course of a cognitive phenomenon (a dynamic, representational activity) may be plotted on a graph whose axes are the PD ratio of the cognitive trails and the degree of stabilization of the cognitive trails. Let us suppose that an activity starts out with low PD ratio and low stabilization. As the field starts to become structured – the creatures start to find their way around a landscape (as the theorist would say) – PD ratio will increase. A network of cognitive trails is temporarily established, and this provides for the possibility of stabilization. Both stabilization and PD ratio continue to increase, until the work concentrates almost entirely on the stabilization of trails that are in place. However, once a network of trails is tightly stabilized it becomes less flexible, and as the nature of the field of activity changes over time, PD ratio will start to decrease as stabilization increases. Further improvement in

way-finding will then require that a destabilized region of cognitive trails be established for a period of time in order to allow PD ratio to increase again. In other words, *virtuous* representational activity is the effective trade-off of the relative merits and demerits of PD ratio and stabilization. Virtuous activity may itself be represented as a figure, a shape, in the two-dimensional space of the PD ratio/stabilization graph. It is not hard to see that the virtuous form of representational activity has the shape of a spiral.” (Cussins, 1993, p. 249-250)



**Figure 3.** The spiral of virtuous representational activity (Cussins, 1993, p. 250).

Cussins’ theory of cognitive trails depicts change as construction and maintenance of multiple intersecting trails across an unexplored field or terrain, gradually leading to a stable conceptualization of the terrain, and subsequently again to destabilization. The terrain may be understood as a zone of proximal development for the actors who enter it. The terrain, or the zone, is both material and mental. Importantly, the theory of cognitive trails has no built-in linear directionality, closure or finalism (‘refinement’, ‘endpoints’) in it. The trails take multiple directions, there is no untouchable predetermined goal, and the contents of the emerging concept are open.

What the theory of cognitive trails fails to discuss is the fact that there are no empty terrains to begin with. They are already inhabited by objects, actors and activities, and the networks those have established are often very stable and resistant. In this light, the crucial question is: What makes change possible in the first place?

This is where activity theory builds on the concept of contradictions. No terrain of activity, no matter how stable and resistant, is free on inner contradictions. Contradictions are not the same as problems or conflicts. Contradictions are historically accumulating structural tensions within and between activity systems. An activity system is constantly working through tensions and contradictions within and between its elements. Contradictions manifest themselves in disturbances and innovative solutions. In this sense, an activity system is a virtual disturbance- and innovation-producing machine.

The primary contradiction of activities in capitalism is that between the use value and exchange value of commodities. This primary contradiction pervades all elements of our activity systems. The school-going activity of students may serve as an illustration. The primary contradiction, the dual nature of use value and exchange value, can be found by focusing on any of the elements of the student’s activity. A student works on textbooks. But textbooks are not utilized only for learning useful knowledge - they are also, and often primarily, studied in order to get grades, to maximize the exchange value of the student as a marketable product.

Activities are open systems. When an activity system adopts a new element from the outside (for example, a new technology or a new object), it often leads to an aggravated secondary contradiction where some old element (for example, the rules or the division of labor) collides with the new one. Such contradictions generate disturbances and conflicts but also innovative attempts to change the activity, making the zone of proximal development an invisible battleground. The stiff rules lagging behind and thwarting possibilities

opened up by advanced new instruments are a common example. A typical secondary contradiction in the activity of school-going may be, for instance, triggered by the introduction of computers and Internet into the students' work. Internet opens up a huge range of interesting and entertaining objects that potentially jeopardize the school's control over students' attention and effort in classrooms.

Contradictions are not just inevitable features of activity. They are "the principle of its self-movement and (...) the form in which the development is cast" (Ilyenkov, 1977, p. 330). This means that new qualitative stages and forms of activity emerge as solutions to the contradictions of the preceding stage of form. This in turn takes place in the form of 'invisible breakthroughs', innovations from below.

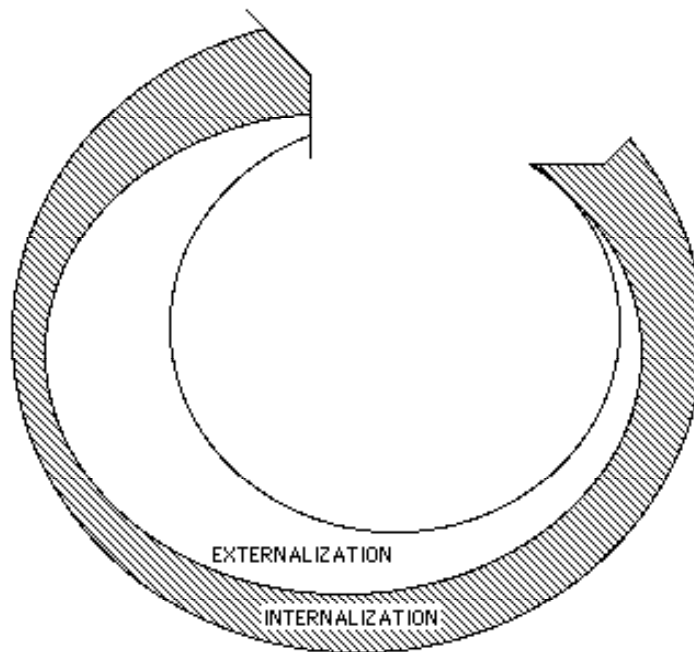
"In reality it always happens that a phenomenon which later becomes universal originally emerges as an individual, particular, specific phenomenon, as an exception from the rule. It cannot actually emerge in any other way. Otherwise history would have a rather mysterious form.

Thus, any new improvement of labour, every new mode of man's action in production, before becoming generally accepted and recognised, first emerge as a certain deviation from previously accepted and codified norms. Having emerged as an *individual exception* from the rule in the labour of one or several men, the new form is then taken over by others, becoming in time a new *universal norm*. If the new norm did not originally appear in this exact manner, it would never become a really universal form, but would exist merely in fantasy, in wishful thinking." (Ilyenkov, 1982, p. 83-84)

Breaking away from a pre-existing trail or terrain requires expansive agency. This can be achieved by employing external cultural artifacts that are invested with meaning and thus become powerful mediating signs that enable the human being to control his or her behavior from the outside. This is the mechanism of double stimulation. It is often interpreted merely as a way to enhance performance in specific tasks of learning and problem solving. Such a technical interpretation neglects the developmental significance of double stimulation as essentially a mechanism of building novel concepts, agency and will.

Contradictions trigger change and development, but they alone do not explain why some change efforts die while some others evolve into long, sustained transformations which seem to energize themselves. To grasp this phenomenon, activity theorists have looked into dialectical oscillations within developmental transformations. Vygotsky and Leont'ev pointed out a foundational oscillation between internalization of given cultural meanings and externalization of novel ideas or solutions. (2) This is echoed in Tomasello's (1999) description of cultural learning as oscillation between lengthy periods of imitation and intense phases of innovation that lead to externalized, artifact-supported new patterns of action. I prefer to connect this oscillation to the idea of expansive cycles of learning and development (Figure 4).

The expansive cycle depicted in Figure 4 begins with an almost exclusive emphasis on internalization, for example in the sense of socializing and training the novices to become competent members of the activity as it is routinely carried out. Creative externalization occurs first in the form of discrete individual deviations and mundane innovations. As the disruptions and contradictions of the activity become more demanding, internalization increasingly takes the form of defense on the one hand and critical self-reflection on the other hand. Externalization, search for novel solutions, increases and reaches its peak when a new model and germ-cell concept for the activity is designed and implemented. As the new model is stabilized, internalization of its inherent ways and means again becomes the dominant form of learning and development.



**Figure 4.** Oscillation of internalization and externalization in an expansive learning cycle (Engeström, 1999a, p. 34).

It is not too difficult to see similarities between this account and the spiral of virtuous representational activity of Cussins presented above.

In fact, Cussins proposes another, complementary oscillation, namely that between stabilization and destabilization. It is closely related, while not identical, to both internalization/externalization and imitation/innovation. Several interesting attempts have been made to capture the specificity of the actions by means of which human beings construct new meanings and change their own patterns of activity. These include the notions of scaffolding (Davis & Miyake, 2004), bridging (Granott, Fischer & Parziale, 2002), and blending (Fauconnier & Turner, 2003).

4) *Generalization.* Activity theory takes seriously the participants, the local practitioners. It does not assume that the researcher has a magic formula with which he or she can objectively decipher how the participants understand and judge the unfolding events. Instead, the practitioners themselves are asked to look at, comment on and make sense of the researcher's initial data and provisional analysis. Ever since our initial workplace studies in the early 1980s (e.g., Engeström & Engeström, 1986), we have routinely shown work sequences we have videotaped to the workers themselves and asked them to interpret the events. The ensuing dialogue itself becomes a new layer of data that gives voice to the practitioners' interpretations (Engeström, 1999b).

In such a dialogical and longitudinal relationship, the researcher is interested in the practical, material generalization of novel solutions and developmental breakthroughs. These solutions are articulated with the help of new concepts and models. For the researcher, such new concepts and models become findings that can acquire significant theoretical import. For the practitioners, those concepts and models are tools that either die out or stabilize and spread. In the latter case, they are typically borrowed and hybridized with other concepts and conditions in other activity systems. This complex process of generalization through practice-bound hybridization represents an alternative way to look at generalizability and is itself a fascinating object of study.

diSessa and Cobb (2004) argue convincingly that design experiments should aim at the generation of 'ontological innovations', understood as new explanatory constructs or theoretical categories. They offer two examples, namely the concepts of 'meta-representational competence' and 'socio-mathematical norms'. Both ideas emerged initially as unplanned and surprising byproducts of design experiments, to be subsequently elaborated and extended into generative concepts. Although diSessa and Cobb do not discuss their ontological innovations in terms of generalization, they may be interpreted as examples of processes of theoretical generalization.

Vygotsky was keenly aware of the need for genuine theoretical generalizations. He pointed out that Marx analyzed the 'cell' of capitalist society in the form of the commodity value: "He discerns the structure of the whole social order and all economical formations in this cell" (Vygotsky, 1997a, p. 320). Vygotsky

continued citing Engels (1925/1978, p. 497) for whom such a cell “represents the process in a pure, independent and undistorted form”. In the first chapter of *Thinking and Speech* (1987), Vygotsky presented the famous contrast between analysis into elements and analysis into units.

“In contrast to the term ‘element’, the term ‘unit’ designates a product of analysis that possesses *all the basic characteristics of the whole*. The unit is a vital and irreducible part of the whole.” (Vygotsky, 1987b, p. 46)

A genuine theoretical generalization is thus based on a ‘cell’ that represents a complex system in a simple, ‘pure’ form. Such a cell retains all the basic characteristics and relationships of the whole system. It is also an ever-present, common part of the whole.

Davydov (1990) subsequently developed these insights into a fully elaborated theory of generalization. In Davydov’s analysis, theoretical generalization is a multi-step process in which an abstract germ cell is first constructed by means of transforming the initial situation experimentally and analytically, and then modeling the emerging idea. The cell is studied by testing and transforming the model. Subsequently, the cell is used to construct increasingly complex extensions and applications, as well as to reflect on and control the very process of generalization. The process leads to rich, continuously expanding living system, the conceptually mastered concrete.

Aspects of theoretical generalization may be identified in the ontological innovations presented by diSessa and Cobb (2004). Initial ingredients of the concepts of ‘meta-representational competence’ and ‘socio-mathematical norms’ were first found when sensory-concrete situations were transformed in design experiments. The elucidation of these concepts eventually led to the construction of expanding systems of applications and extensions. What seems to be missing in the accounts of these ontological innovations is the modeling of the germ cells. In other words, the simple and pure anatomy of these concepts is not shown.

## The Change Laboratory

In the mid-1990s, researchers in the Center for Activity Theory and Developmental Work Research at University of Helsinki developed a new intervention toolkit under the generic name of *Change Laboratory* (Engeström & al., 1996, Engeström, 2007b). Variations of this toolkit have been used in a large number of intervention studies in settings ranging from post offices and factories to schools, hospitals and newsrooms. The Change Laboratory serves as a microcosm in which potential new ways of working can be experienced and experimented with (Engeström, 1987, p. 277-278).

A Change Laboratory is typically conducted in an activity system that is facing a major transformation. This is often a relatively independent pilot unit in a large organization. Working practitioners and managers of the unit, together with a small group of interventionist-researchers, conduct five to ten successive Change Laboratory sessions, often with follow-up sessions after some months. When feasible, also customers or patients are invited to join Change Laboratory sessions in which their particular cases are analyzed in detail. Change Laboratories are also conducted as boundary crossing laboratories with representatives from two or more activity systems engaged in collaboration or partnership.

The Change Laboratory is built on ethnographic data from the activity setting in which it is conducted. Critical incidents, troubles and problems in the work practice are recorded and brought into Change Laboratory sessions to serve as *first stimuli*. This ‘mirror material’ is used to stimulate involvement, analysis and collaborative design efforts among the participants.

To facilitate analysis and resolution of the problems, interventionists typically introduce conceptual tools such as the triangular models of activity systems (see Figure 1) as *second stimulus*. Commonly the conceptual models offered by the interventionists are replaced or combined with mediating conceptualizations or models formulated by the participants.

The participants are challenged to use the mediating second stimulus as an instrument in the design of a *new concept* for the activity they are trying to transform. Implementation of the designed new solution is usually initiated while the Change Laboratory sessions are still running, in the form of pilot experiments. The implementation typically leads to a richer and more articulated concept.

In the analysis and design, the participants are asked to move between the past, the present, and the future. This means that historical origins of the current problems are dug up and modeled, and the ideas toward a future concept are played with in anticipatory simulations such as role play.

The laboratory sessions themselves are videotaped for analysis and use as stimuli for reflection. The procedure allows for the collection of rich longitudinal data on the actions and interactions involved in cycles of expansive learning (Engeström, 2001).

## Recapturing the object in a hospital surgery unit

Earlier in this paper I pointed out that there are three main differences between linear interventions (including design experiments) and formative interventions. The three may be named 1) fixed goals vs. poorly understood contradictory object as starting point, 2) reception vs. double stimulation as process, and 3) standardized solution packages vs. intermediate conceptual tools as outcome. In the following, I will examine data from a recent Change Laboratory intervention in the light of these three crucial differences.

The Central Surgery Unit of the University Hospital of the City of Oulu in northern Finland invited my research team to work with their staff in the fall of 2006. The unit was suffering from shortages of qualified staff and excessive workloads which led to temporary closures of some of the operating theaters. This, in turn, led to longer waiting times and queues for patients, and to negative publicity. The pressures were intensified by the fact that new legislation required all hospitals in the near future to guarantee patients access to care with short waiting times, which meant that the hospitals would have to take determined administrative action to cut their queues. The management of the Central Surgery Unit was aware that their problems had a lengthy history and could probably not be eliminated by quick fixes. Our task was to conduct a collective process in which the problems would be diagnosed and holistic long-term solution would be sought.

We conducted six two-hour Change Laboratory sessions in the fall of 2006. We also conducted follow-up sessions in 2007 and 2008. The participants of the sessions were selected to represent the whole range of practitioners working in the unit, from the head doctor of the unit to surgeons, anesthesiologists, nurses, a porter and a secretary. We videotaped the sessions and transcribed the discussions.

In the first session, conducted on September 28, 2006, there were 21 participants: 4 surgeons, 4 anesthesiologists (one whom served as the operations manager of the whole unit), 3 surgical nurses, 4 anesthesia nurses (including the head nurse of the unit), one administrator, one secretary, one porter, and 3 members of the research group. (3) In the two-hour session, there were 402 conversational turns. In the following analysis, I will only use the transcript of the first session as data. In the following, I will scrutinize the data to characterize the starting point, the process, and the provisional outcome of the Change Laboratory intervention.

We started the first session by presenting a set of videotaped interviews in which practitioners expressed their concern with the shortages of staff, excessive workloads, long waiting lists, and closures of operating theaters. This 'mirror material' triggered a rich discussion.

1) *Starting point: Loss of the object.* Much of the discussion in the first session circled around experiences of loss of meaning in the work. These experiences were expressed in different thematic streams. The first one focused on the vicious circle of necessary temporary closings of operating theaters and growing patient waiting lists.

048     *Researcher 1:* (...) is it some kind of a problem, that they have to stay closed, the operating theaters?

049     *Anesthesiologist, operation manager:* It's a problem in the sense that there are patients in the waiting list who need the operations, and there is a lot of pressure on the other side that they must get treated. This can be seen in the public. So you are between a rock and a hard place all the time. Which gives you the feeling that here we are constantly failing, even though we work harder than ever, we are bad all the same because we cannot get the waiting list to move.

The formulation "between a rock and a hard place all the time" is a clearcut expression of what Bateson (1972) called a double bind: "we are bad all the same."

The second stream focused on the difficult or 'bad' patient material in the Central Surgery Unit.

087     *Head anesthesia nurse:* If I say on behalf of the anaesthesia nurses so we really have three different anaesthesia units. So we have this ours, the central clinic, and then short-term day surgery. If we compare the working conditions, patient material, working time, so this is quite different. So we work three shifts, we have the emergency service duty. The patient material...

088     *Researcher 1:* (...) please say that again, wind back.

089     *Head anesthesia nurse:* Well. So the patient material is surely the worst of all. And the operations are the biggest, and the operations are the heaviest...  
(...)

092     *Head anesthesia nurse:* Yes. And quite the same pay. And then the patient material, that I already mentioned. It is worse and it will remain worse because now the private sector combs out the easiest patients and operates on them there at a different time when we get all the worse patients. Then a big problem for us is the postoperative care, the immediate

postoperative care, in other words the recovery room which gets not only all our patients, but patients from the central clinic, and the emergency service, internal medicine patients, intensive care patients...

Here the discourse shifted from 'us' (the practitioners) to 'them' (the patients), that is, from the subject to the object. But the talk about 'bad patient material' displayed a reified object. Patients have become burdensome anonymous material. This was quickly countered by a surgeon who took up the other side of patients.

111 *Surgeon 3:* I find quite contradictory when you think about the people who leave and say it is because the job is so hard. I just feel that in this job, if you do the real work with the patient, however hard it may be, so it is just that that is rewarding. On the other hand you are ready to take on the challenge if you have the education for it. That the leavings, so they surely come from something else than that you have a difficult case in front of you, and that is what you battle with. So that when you leave to go home, you feel relieved, a nice feeling that I bloody did it! That you don't think this is so hard that I will leave this job.

112 *Researcher 1:* So here we have a comment that the fact that the patient material is becoming harder is not necessarily directly the reason.

113 *Senior anesthesiologist:* But it can be a reason in the way that if you experience the hardness and not the reward at all. Nobody ever thanks you, the work remains unfinished, you leave, and everything is unfinished, your shift ends and you leave; you never see the reward for the hard work, so that... Surely one thing that should be developed is that really you would feel that the work is rewarding. So that you would really feel that you do your job well and that the patients get well and stay alive. And in this way the whole image of the thing should get built.

The senior anesthesiologist (turn 113) actually opened up a third stream in the discussion: "You never see the reward for the hard work." This stream was continued and elaborated on later in the session. The point of this stream was the experience of loss of view and control over the object of one's work.

345 *Senior anesthesiologist:* (...) So you don't get to control your own work so that you would know what they have there-. There we had it again, that I thought a little bit that is it like the chiefs in Asterix that they are afraid of the sky falling down, that they don't know what will be coming in through the door. Why be afraid, why should you be afraid of that? When you have the best professionals, who can keep the person alive, you are afraid of what comes in through the door. But it surely doesn't depend on that, that you would be afraid that you cannot do something to it, but that you don't know in advance about the day, so that takes it away... That's how I see it anyway. Is it such a thing, that there is the scare, that it is the unknown. When you have no idea at all who the patient is going to be. When if you saw on some sort of a list that there we have our orthopaedic patients whom we are going to operate today, these come in through our recovery part and these we are expecting. We will have a look at the monitor screen of how the things will proceed, that we know who is going to come next and then here again.

346 *Senior surgical nurse:* Those we do know. But then the ones, who come in through the central clinic or the ones that come from X-ray or who come in through the emergency room... Who come back from the ward when they cannot cope... Or when it is a patient back there for three days when he should be in the ICU but they don't have the bed. And they won't take him in a corridor place in the in-patient ward and...

347 *Senior anesthesiologist:* So you have no control over your own work and you cannot plan it beforehand, and these are probably what cause the...

348 *Senior surgical nurse:* Yes, in a way they mess up this our own operation.

The starting point of the Change Laboratory intervention in this case consisted of intense articulation of the experienced problem space, or spectrum of trouble. Through its three streams, this articulation transformed the initially given, publicly available problems (shortages of personnel, long waiting lists, closures of operating theaters) into a poorly understood contradictory object.

The first articulated the situation as a double bind. The second stream articulated patients as reified into anonymous 'bad material' (or its flip side, 'challenging material', as in turn 111). This kind of categorization, or empirical abstraction (Davydov, 1990), typically serves as 'stabilization knowledge' that gives a name and a simple pseudo-explanation to a complex and bewildering problem (Engeström, 2007a). The third stream of



articulated the situation as fear for the unknown, as loss of control and overview of what kinds of patients one will encounter during the workday. Taken together, the three streams of this initial articulation may be characterized as *loss of object*.

2) *Process: Alternative organization as second stimulus*. One might assume that the intervention would have proceeded in a moreless straightforward manner to regain and re-identify the lost object. However, there was no such a linear fix available. A second stimulus was needed to open up an indirect avenue toward regaining the object.

In this case, the second stimulus was generated by the participants already in the first session, before we researchers had a chance to suggest our own conceptual tools for the task. This reminds us of Vygotsky's (1999, p. 60) statement quoted earlier: "we do not necessarily have to present to the subject a prepared external means with which we might solve the proposed problem"; we may "wait while he spontaneously applies the auxiliary device and involves some auxiliary system of symbols in the operation."

In this case, the second stimulus, or the mediating conceptual instrument, began to take shape first in relatively vague terms, as complaints about the large size of the unit.

097     *Senior anesthesia nurse*: (...) And the number of staff is large, and really as I said the area of responsibility is big.

098     *Researcher 1*: What's the disadvantage in having large numbers of staff?

099     *Senior anesthesia nurse*: Management is hard and communicating is very difficult, getting the group together is really hard, motivating is hard. And what else should I think? Knowledge management is hard.

The issue was further developed as a need for smaller sub-units.

138     *Head nurse*: (...) If we want to make this education and familiarisation easier we should absolutely divide the people into smaller pools, or specialities, or whatever you call it. Like the surgeons have. In that case it would be easier to take care of.

The actual formulation of the second stimulus was initiated by a surgeon.

193     *Surgeon 1*: Yes. Such a thought that as T (senior anesthesia nurse) said, the Central Surgery Unit is terribly big, and therefore hard to control, that what if we divided it into parts. Orthopedics would get their own department, as would soft tissue surgery, cardio-thoracic surgery, and vascular surgery their own. Into three parts so that each would have their own nurses, their own doctors there, so that we would have smaller units, easier to manage, better to build such own identity for each and everyone and easier to recruit new people. Would that be more functional?

194     *Researcher 1*: Have you seen similar models in Finland?

195     *Surgeon 1*: Yes in Tampere.

196     *Researcher 1*: Does it work?

197     *Surgeon 1*: It works very well.

The idea of dividing the unit into three or four smaller independent sub-units was received with both enthusiasm and reservations. The new organizational model evoked considerations of complications. However, these were taken as issues to be resolved by joint design efforts.

204     *Senior anesthesia nurse*: And then such a thing that in surgery and instruments it works and it is good that there are the doctors who are responsible in each area of expertise. But then this kind of a general, that the specializing doctors and anesthesia nurses, so for them the rotation is absolutely vital because we have the emergency duty and such. So in a way in the anesthesia side you have to have flexibility in building the system.

205     *Researcher1*: If it would be three separate units, then you should be able to move across them.

206     *Senior anesthesia nurse*: Yes.

207     *Researcher 1*: Exactly. That it brings along its own difficulties right from the start.

208     *Operations manager*: And it also leads to the fact that the emergency duty should be partly made separate which we already have in the plans, on the other hand, that we will have the soft tissue and hard tissue emergency services.

209     *Researcher 1*: Do you mean that in a way there would be a fourth unit, the emergency service unit?

210     *Operations manager:* No, I mean that if we were to be divided into these smaller units, so in a way the units would take care of their own emergency service. In any case we have three to four teams there, but they would represent the knowledge in the different areas of expertise.

The construction of the second stimulus, namely the alternative organizational structure based on smaller sub-units, was a long and laborious process. It went on practically all through the six Change Laboratory sessions, resulting in a detailed organization model. This new model was subsequently implemented in the Central Surgery Unit. The relevance of double stimulation in long-term processes was not alien to Vygotsky. He pointed out that the principle applies even to “a continuing real-life problem that ends only at death and whose resolution may be interrupted thousands of times” (Vygotsky, 1997b, p. 213).

Where did the idea for the second stimulus come from? An exchange between four physicians indicates that the idea had been around for quite a while.

216     *Surgeon 2:* I feel that we are mixing things a little bit. The things that V (surgeon 1) said about Tampere, they are organizationally different units. And I think that what you said about dividing the functions is nothing but regrouping the functions here.

217     *Operations manager:* They could have their own organizations, we don't...

218     *Surgeon 1:* This is a step into that direction.

219     *Surgeon 2:* So it is a really long step. They would have to have their own ward. They would have their own financial management, own management...

220     *Surgeon 1:* We have talked about this for a long time.

221     *Surgeon 3:* It is just like V (surgeon 1) said, and at least I am in favour warmly and vote for V (surgeon 1) even in the council elections (laughter). This is a work policy issue. The knowledge base in surgery is so huge. In addition I know that this has been talked about in several meetings and talks. It is exactly anesthesiologists, anesthesia nurses and instrument nurses who bring this up a lot; they do so much that the control of all this should be, it is an impossible goal. It is a fact that going through all the theaters is not enough to take control of them. You have to take realism by the hand and do your own specializations.

Vygotsky talked about the second stimulus as ‘neutral stimulus’ which is filled with significance. In the simple examples of tying a knot or using the position of the hands of the clock as a prompt, the second stimuli look indeed neutral. But no object is neutral. By virtue of being recognized as an object it already is invested with a meaning and certain affordances which have a history. In this sense, the notion of ‘neutral’ is unfortunate and should be replaced with ‘ambiguous’. The initially ambiguous artifact – in our case, the alternative organizational model – is gradually filled with specific contents and demands for action. The crucial issue is that the second stimulus is a mediating artifact, not the solution in itself. Adopting a new organizational model is not an automatic solution to the loss of the object. It has to be implemented as an instrument to solve the problem, which in turn means that the new solution, the new concept, only emerges as the second stimulus is put into action.

3) *Outcome: New concept of responsibility?* From its very inception, the mediating second stimulus (the alternative organizational model) was inherently ambiguous in that it implied two possible and partly conflicting outcomes. The first one was specialization.

228     *Surgeon 4:* (...) the central operating ward it is a big ward, and a large part of it, apparently all procedures, especially in the instrument side, they go around the whole system all the time. And this leads to the fact that we have these specializations, where e.g. in the instrument side you are given on the table five containers of nuts and bolts, and it always takes about 2—3 months to learn all those nuts and bolts. And the training restarts every half year or a year, and the first 2—3 months slow down the operations and efficiency. And then when we get to the point where it starts to run smoothly, we start the same round again. It is continuing learning which affects the normal operations-, it is inefficient. It always slows down the action and it is not rational for the fact that when the person has been there for six months or a year s/he then goes elsewhere and comes back again after five years, s/he has had time to forget everything several times over. In addition it is completely useless. And not everybody needs to know all of these things.

(...)

235     *Surgery nurse 1:* We could also say that on the contrary it is also so that when you have an experienced nurse s/he sometimes oversees the inexperienced surgeon that s/he does things right. But about the emergency service time, I find it contradictory that we have doctors

in all the specialties, own doctors who do the emergency duty – as back-up – but we don't have in the specialty in question a back-up emergency nurse who knows all the work in the area exactly, who we could ask at least by telephone. And then we do there, we sort of sail through from one day to the next.

The other possible outcome of the new organizational model was characterized as a new kind of shared responsibility.

202 *Senior surgery nurse:* I feel that taking the responsibility would perhaps be-, or should I say, that there would be more people taking the responsibility when we would have such a smaller system. That now it is easy to throw everything to P (the operations manager) and maybe some little goes to T (senior anesthesia nurse), too.  
(...)

266 *Operations manager:* we live at the point of whether we drown or not. Each model or thing that we do puts more water into the boat. And then if we were divided into smaller units, then perhaps the ability to see the whole would grow among this group. Sometimes I feel that everyone just thinks that they aren't interested, I do exactly what I have been told to do and I don't care how this thing gets done as a whole. And seeing the whole is then left with the small group in the control room who try to fight the big current.

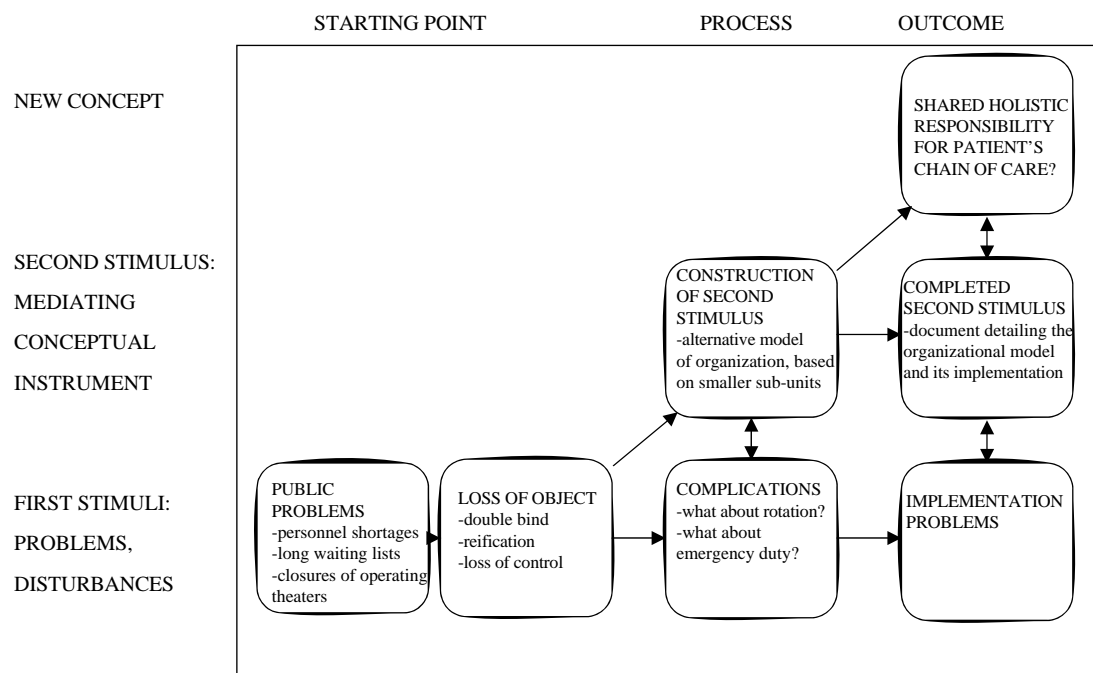
In a way, specialization and shared responsibility were opposite images of the outcome. Specialization implies focusing on a *narrower* field, with more depth. The new responsibility called for in turns 202 and 266 envisioned collective focus on a *wider* field: "to see the whole." The need for such a wider or more holistic perspective was formulated in terms of the patient's "whole chain of care."

354 *Head nurse:* And if you think of the whole chain of care, then we have a big bucket into which everyone pours, and out of the bucket comes a thin hose which leads to the operating ward, and from thereon in its time to the in-patient wards. There are a number of narrow passages and at the end there is a plug. This is because of the fact that the patients for several reasons cannot go home or to further care. And the ward capacity for in-patients is not enough for this kind of patient numbers. And there too, when we talked that this is a kind of a spiral that in my mind in this patient flow it is such that when the patients wait for days, even four days, for their emergency operations in the ward, so they cannot get into the operating theatre, and therefore the patients who have already been operated on and are in the recovery ward cannot get into the in-patient wards. This is if you think in caricature. Or then when the patient leaves the ward to go to the operating theatre, and e.g. it has been planned that there will be 1—2 days in the intensive care unit then his or her place in the in-patient ward will be filled not with a patient who will be there for the one or two days that the patient is away, but similar kinds of patients who need long-term care. So when we want to transfer the patient from either the recovery ward or the ICU to the ward, then the ward announces that we cannot take him or her as we have no places as we took in 10 new patients today. Emergencies, or elective patients too.

It is too early to say to which extent and in which ways this new, more holistic responsibility for the patient's chain of care may have emerged as an outcome of the formative intervention in this case.

## **The layers of formative intervention**

The analysis presented above may be summarized with the help of Figure 5. The diagram shows the layered character of the formative intervention.



**Figure 5.** The layered character of a formative intervention.

As the intervention progresses from the starting point to the construction of the second stimulus, the first stimuli (the problems to be solved) also gain a new layer, appearing as complications engendered by the second stimulus itself. Similarly, as the intervention moves to the phase of implementing the second stimulus in practice and thus generating a new concept, the second stimulus as well as the first stimuli also gain a new layer. Although this goes beyond the data analyzed for the present paper, it may be noted that the second stimulus took the shape of a sizeable document, entirely written by the practitioners themselves, that detailed the new organizational model and guidelines for its implementation. The first stimuli took the shape of pressing practical problems engendered by the implementation of the new organizational model.

## Formative interventions in schools

Formative interventions of the Change Laboratory type have mainly been conducted in workplaces. Can they also be used in schools? After all, design experiments are firmly rooted in schools. If formative interventions are inapplicable in school settings, perhaps the two methods just cannot be compared and my criticism of design experiments is misplaced.

In fact, Change Laboratories have been conducted in schools. My research group ran a series of successful Change Laboratories in a middle school in Helsinki (Engeström, Engeström & Suntio, 2002a, 2002b). The laboratory sessions were attended by the teachers of the school. This leaves open the question whether students can be central participants in a Change Laboratory.

A first step to this direction was taken by Marianne Teräs who organized and analyzed a version of the Change Laboratory, called Culture Laboratory, with immigrant students and teachers of Helsinki City College of Social and Health Care (Teräs, 2007). Implementing formative interventions with students of primary and middle schools will, however, be a new challenge.

Why should such an attempt be made? How could young students contribute to the development of their schooling? Motivation is foundational for meaningful school work. Yet educational institutions are notoriously bad at understanding and nurturing students' motivation. Our main motivational instruments are still grades, used to reward and punish. Students are the best experts in their own motivation, including critical awareness of factors that prevent or distort their motivation to learn. Formative interventions can give voice to this expertise and open up new ways to build motivation, or more appropriately, to turn motivation into agency.

## Endnotes

- (1) There may be design experiment studies that overcome the limitations described in the three points of contrast. This paper is not meant to be a comprehensive assessment of work done within the paradigm of design experiments. Thus, my critique is necessarily somewhat simplified and treats design research as a totality without much nuance.
- (2) Another important developmental oscillation is that between socially oriented and object-oriented activity, presented by El'konin (1977) as the explanatory principle behind shifts in leading activity in the development of children and adolescents. Space does not permit me to discuss this oscillation more fully here.

- (3) The research group consisted of Yrjö Engeström (PI), Hannele Kerosuo and Anu Kajamaa. Also Philippe Lorino and Annalisa Sannino have participated in data analysis of the transcript of the first Change Laboratory session. I am grateful for their insights. However, I am alone responsible for the interpretations included in the preliminary analysis presented in this paper. This analysis is very much work in progress.

## References

- Ausubel, D. P. (2000). *The acquisition and retention of knowledge: A cognitive view*. Dordrecht: Kluwer.
- Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher*, 32, 21-24.
- Barab, S. A. & Kirshner, D. (2001). Guest editors' introduction: Rethinking methodology in the learning sciences. *The Journal of the Learning Sciences*, 10, 5-15.
- Bateson, G. (1972). *Steps to an ecology of mind*. New York: Ballantine Books.
- Brown, A. L. (1992). Design experiments: Theoretical, and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, 2, 141-168.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R. & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32, 9-13.
- Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology*. Berlin: Springer.
- Collins, A., Joseph, D. & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, 13, 15-42.
- Cussins, A. (1992). Content, embodiment and objectivity: The theory of cognitive trails. *Mind*, 101, 651-688.
- Cussins, A. (1993). Nonconceptual content and the elimination of misconceived composites! *Mind & Language*, 8, 234-252.
- Davis, E. A. & Miyake, N. (Eds.) (2004). *Special issue on scaffolding*. *The Journal of the Learning Sciences*, 13(3).
- Davydov, V. V. (1990). *Types of generalization in instruction: Logical and psychological problems in the structuring of school curricula*. Reston: National Council of Teachers of Mathematics.
- diSessa, A. A. & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *The Journal of the Learning Sciences*, 13, 77-103.
- El'konin, D. B. (1977). Toward the problem of stages in the mental development of the child. In M. Cole (Ed.), *Soviet developmental psychology*. White Plains: Sharpe.
- Engels, F. (1925/1978). *Dialektik der Natur*. Berlin: Dietz.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsultit.
- Engeström, Y. (1989). The cultural-historical theory of activity and the study of political repression. *International Journal of Mental Health*, 17:4, 29-41.
- Engeström, Y. (1999a). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen & R-L. Punamäki (Eds.), *Perspectives on activity theory*. Cambridge: Cambridge University Press.
- Engeström, Y. (1999b). Expansive visibilization of work: An activity-theoretical perspective. *Computer Supported Cooperative Work*, 8, 63-93.
- Engeström, Y. (2000). From individual action to collective activity and back: Developmental work research as an interventionist methodology. In P. Luff, J. Hindmarsh & C. Heath (Eds.), *Workplace studies*. Cambridge: Cambridge University Press.
- Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133-156.
- Engeström, Y. (2004a). Managing as argumentative history-making. In R. Boland (Ed.), *Managing as designing*. Stanford: Stanford University Press. p. 97-101.
- Engeström, Y. (2004b). New forms of learning in co-configuration work. *Journal of Workplace Learning*, 16, 11-21.
- Engeström, Y. (2005). *Developmental work research: Expanding activity theory in practice*. Berlin: Lehmanns Media.
- Engeström, Y. (2007a). From stabilization knowledge to possibility knowledge in organizational learning. *Management Learning*, 38, 271-275.
- Engeström, Y. (2007b). Putting Vygotsky to work: The Change Laboratory as an application of double stimulation. In H. Daniels, M. Cole & J. V. Wertsch (Eds.), *The Cambridge companion to Vygotsky*. Cambridge: Cambridge University Press.
- Engeström, Y. & Engeström, R. (1986). Developmental work research: The approach and the application of cleaning work. *Nordisk Pedagogik*, 6, 2-15.
- Engeström, Y., Lompscher, J. & Rückriem, G. (Eds.) (2005). *Putting activity theory to work: Contributions from developmental work research*. Berlin: Lehmanns Media.

- Engeström, Y., Engeström, R. & Kerosuo, H. (2003). The discursive construction of collaborative care. *Applied Linguistics*, 24, 286-315.
- Engeström, Y., Engeström, R. & Suntio, A. (2002a). Can a school community learn to master its own future? An activity-theoretical study of expansive learning among middle school teachers. In G. Wells & G. Claxton (Eds.) *Learning for life in the 21<sup>st</sup> century: Sociocultural perspectives on the future of education*. London: Blackwell.
- Engeström, Y., Engeström, R. & Suntio, A. (2002b). From paralyzing myths to expansive action: Building computer-supported knowledge work into the curriculum from below. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community*. Hillsdale: Lawrence Erlbaum.
- Eskola, A. (1999). Laws, logics, and human activity. In Y. Engeström, R. Miettinen & R-L. Punamäki (Eds.), *Perspectives on activity theory*. Cambridge: Cambridge University Press.
- Fauconnier, G. & Turner, M. (2003). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.
- Granott, N., Fischer, K. W. & Parziale, J. (2002). Bridging to the unknown: A transition mechanism in learning and development. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning*. Cambridge: Cambridge University Press.
- Haenen, J. (1995). *Pjotr Gal'perin: Psychologist in Vygotsky's footsteps*. Hauppauge: Nova Science Publishers.
- von Hippel, E. & Tyre, M. J. (1995). How learning by doing is done: Problem identification in novel process equipment. *Research Policy*, 24, 1-12
- Holt, J. (1964). *How children fail*. New York: Dell.
- Il'enkov, E. V. (1977). *Dialectical logic: Essays in its history and theory*. Moscow: Progress.
- Il'enkov, E. V. (1982). *The dialectics of the abstract and the concrete in Marx's 'Capital'*. Moscow: Progress.
- Kelly, A. (2004). Design research in education: Yes, but is it methodological? *The Journal of the Learning Sciences*, 13, 115-128.
- Leont'ev, A. N. (1932). The development of voluntary attention in the child. *Journal of Genetic Psychology*, 40, 52-81.
- Long, N. (2001). *Development sociology: Actor perspectives*. London: Routledge.
- Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, 33, 3-11.
- Melucci, A. (1996). *Challenging codes: Collective action in the information age*. Cambridge: Cambridge University Press.
- Midgley, G. (2000). *Systemic intervention: Philosophy, methodology, and practice*. New York: Kluwer.
- Olson, D. R. (2004). The triumph of hope over experience in the search for "what works": A response to Slavin. *Educational Researcher*, 33, 24-26.
- Talyzina, N. F. (1981). *The psychology of learning: Theories of learning and programmed instruction*. Moscow: Progress.
- Teräs, M. (2007). *Intercultural learning and hybridity in the Culture Laboratory*. Helsinki: University of Helsinki, Department of Education.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press.
- van der Veer, R. & Valsiner, J. (1991). *Understanding Vygotsky: A quest for synthesis*. Oxford: Blackwell.
- Victor, B. & Boynton, A. C. (1998). *Invented here: Maximizing your organization's internal growth and profitability*. Boston: Harvard Business School Press.
- Vygotsky, L. S. (1978). *Mind in society: The psychology of higher mental functions*. Cambridge: Harvard University Press.
- Vygotsky, L. S. (1987a). Lectures on psychology. In *The collected works of L. S. Vygotsky. Vol. 1. Problems of general psychology*. New York: Plenum.
- Vygotsky, L. S. (1987b). Thinking and speech. In *The collected works of L. S. Vygotsky. Vol. 1. Problems of general psychology*. New York: Plenum.
- Vygotsky, L. S. (1997a). The historical meaning of the crisis in psychology: A methodological investigation. In *The collected works of L. S. Vygotsky. Vol. 3. Problems of the theory and history of psychology*. New York: Plenum.
- Vygotsky, L. S. (1997b). The history of the development of higher mental functions. In *The collected works of L. S. Vygotsky. Vol. 4. The history of the development of higher mental functions*. New York: Plenum.
- Vygotsky, L. S. (1997c). The instrumental method in psychology. In *The collected works of L. S. Vygotsky. Vol. 3. Problems of the theory and history of psychology*. New York: Plenum.
- Vygotsky, L. S. (1999). Tool and sign in the development of the child. In *The collected works of L. S. Vygotsky. Vol. 6. Scientific legacy*. New York: Kluwer/Plenum.
- Zinchenko, V. P. (1985). Vygotsky's ideas about units for the analysis of mind. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives*. Cambridge: Cambridge University Press.