THE FUTURE OF ACTIVITY THEORY: A ROUGH DRAFT*

Yrjö Engeström
University of Helsinki

INTRODUCTION

Recently Aaro Toomela has published two papers in which he critically assesses the contents and potential of activity theory. The first one is titled 'Activity theory is a dead end for cultural-historical psychology' (Toomela, 2000), the second one 'Activity theory is a dead end for methodological thinking in cultural psychology too' (Toomela, 2008). According to Toomela, there are five fatal faults in activity theory:

1. It relies on unidirectional instead of a dialectical view of culture-individual relationships.
2. It focuses on analyses of activities without taking into account the individual involved in the activity at the same time.
3. It underestimates the role of signs and the importance of focusing on sign meaning.
4. It approaches mind fragmentally, without understanding the holistic nature of mind.
5. It is fundamentally a developmental and therefore not appropriate for understanding emerging phenomena, including mind.

Toomela presented this condemning assessment first in 2000. If activity theory was indeed a dead end in 2000, one would have expected signs of death or at least withering away in the scientific community. In fact, the opposite has happened, as evidenced in Figure 1, taken from the review article of Roth and Lee (2007, p. 188). The figure shows a rather dramatic growth between 2000 and 2005 in the frequency of journal citations to key activity-theoretical terms and texts, such as the term 'activity theory', the two books of Leont’ev available in English (although sold out long ago), my book Learning by Expanding, and an

---

article from 1993 by Cole and Engeström. There is good reason to assume that this growth has continued strongly also after 2005.

Figure 1. Four indicators of the increasing interest shown in cultural-historical activity theory over the past three decades, based on citation frequencies in the Institute for Scientific Information’s citation database (Roth & Lee, 2007, p. 188)

What might explain the discrepancy between Toomela’s death announcements and the factual growth of research based on activity theory? It seems that this is a case of a peculiar form of scientific autism. Toomela, like many other critics, pays no attention to concrete research done on the basis of activity theory. There is not a single reference to recent concrete activity-theoretical studies in either one of his two doomsday articles. In other words, whatever evidence Toomela might have used to ground his critical assessment, this evidence clearly
does not pertain to what activity theorists are doing, how they are applying and developing their theory. Toomela has first created an abstract and frozen image of activity theory – indeed, a dead image – and then used this image to diagnose a set of fatal faults in it. As the diagnosis did not come true, the author found necessary to repeat it eight years later.

There is a lesson to be learned from this. To keep activity theory alive and productive, we need to read each other’s concrete studies, dig into each other’s data and visit each other’s field sites of research and intervention. And we need to keep on overcoming the five fatal faults. After this preamble, I will now lay out my thoughts about the ways forward.

**RUNAWAY OBJECTS**

Activity theory is a theory of object-driven activity. Objects are concerns, they are generators and foci of attention, motivation, effort and meaning. Through their activities people constantly change and create new objects. The new objects are often not intentional products of a single activity but unintended consequences of multiple activities.

The societal relevance and impact of activity theory depend on our ability to grasp the changing character of objects. In the present era, we need to understand and deal with what I have called ‘runaway objects’ (Engeström, 2008).

Runaway objects have the potential to escalate and expand up to a global scale of influence. They are objects that are poorly under anybody’s control and have far-reaching, unexpected effects. Such objects are often monsters: They seem to have a life of their own that threatens our security and safety in many ways. Klein (2007) argues that in present-day capitalism, disasters and shocks are becoming a dominant object, exploited by the economic and political elites to reorganize societal conditions in line with the neoliberal doctrine.

Runaway objects are contested objects that generate opposition and controversy. They can also be powerfully emancipatory objects that open up radically new possibilities of development and well-being. The Linux operating system is a well-known example. There are other, less known but potentially very significant new objects being created.
“In Brazil, the phenomenon is best seen in the million and a half farmers of the Landless People's Movement (MST) who have formed hundreds of cooperatives to reclaim unused land. In Argentina, it is clearest in the movement of ‘recovered companies,’ two hundred bankrupt businesses that have been resuscitated by their workers, who have turned them into democratically run cooperatives. For the cooperatives, there is no fear of facing an economic shock of investors leaving, because the investors have already left.” (Klein, 2007, p. 455)

Contrary to mega-projects (Altshuler & Luberoff, 2003; Flyvbjerg, Bruzelius, & Rothengatter, 2003), most runaway objects do not start out as big and risky. More commonly, they begin as small problems or marginal innovations, which makes their runaway potential difficult to predict and utilize. They often remain dormant, invisible, or unseen for lengthy periods of time, until they burst out into the open in the form of acute crises or breakthroughs.

Leont’ev’s (1978) well-known dictum was that there is no activity without an object. With runaway objects, we may ask: Are there objects without an activity? Whose object is the global warming, for example?

Of course runaway objects do not emerge and exist without human activities. To begin with, they must be identified and named by humans. The very concept of global warming would not exist if experts, researchers, politicians and journalists had not articulated the phenomenon. But which activities take responsibility for such a huge object as global warming?

I have often used the representation depicted in Figure 2 to capture the challenge of constructing a shared object between two or more activity systems.

Figure 2. Two activity systems and a potentially shared object
However, with large runaway objects, the challenge would look more like Figure 3. The are typically numerous activity systems focused on or affiliated with the object. But the object is pervasive and its boundaries are hard to draw. Thus, the positions of the activity systems are ambiguous and they often seem to be subsumed to the object rather than in control of it.

![Figure 3. Large runaway object and activity systems](image)

Big runaway objects tend to be either what used to be regarded as 'natural forces' (diseases, environmental threats) or technological innovations. Such runaway objects are typically seen as objects for relatively exclusive professional expert activities. Patients, victims and users become marginal, or 'rubbish' (Engeström & Blackler, 2005).

The task of activity theory is to recycle rubbish and to turn it into diamonds. This calls attention to being ill, suffering and recovering, rebuilding, using and tinkering as *productive* activities. We need intermediate runaway objects which are less spectacular and more
inviting.

Various social movements try to do just that. Organic farming, Wikipedia, open models of scientific research and publishing are examples. Most such attempts fail or remain marginal. A crucial question is: What gives some objects inherent drawing power?

In a very tentative way, I would suggest some prerequisites. First of all, a benign runaway object must have intrinsic properties that transcend the limits of utilitarian profit motive. In this sense, a benign runaway object is at the boundary between legitimate and illegitimate, sensible and crazy, work and leisure, technology and art. These properties are experienced in acting on and with the object over a long haul, with persistence and patience, oscillating between intensity and withdrawal. The object must yield useful intermediate products, yet remain an incomplete project. The object must be visible, accessible and cumulable - allowing participants to return time and again. There must be effective feedback from and exchange among the participants acting on the object.

In the following sections, I will discuss the five themes of this book in the light of the challenges posed by the emergence of runaway objects.

**UNITS OF ANALYSIS: THIRD GENERATION ACTIVITY THEORY, AND BEYOND**

I have suggested that the evolution of activity theory may be seen in terms of three generations, each building on its own version of the unit of analysis (Engeström, 1996). The first generation built on Vygotsky’s notion of mediated action. The second generation built on Leont’ev’s notion of activity system. The third generation, emerging in the past 15 years or so, built on the idea of multiple interacting activity systems focused on a partially shared object.

Scholars such as Frank Blackler and his colleagues (e.g., Blackler & McDonald, 2000) have taken up the weak treatment of the issue of power in activity theory. It is indeed not easy to depict and analyze hierarchical power relations within a single activity system. Third generation activity theory may open up new possibilities. In an organization, managing is usually best seen as an activity system of its own, relatively independent of the activity systems of primary productive work. A useful minimal unit of analysis might in some cases
look like the diagram in Figure 4. In the diagram, the relationship between the activities of management and work, specifically the flow of rules from management to work units, is opened up for scrutiny. Yet these two activity systems and their takes on the potentially shared object are looked at in relation to the activity system of the client. Examination of the horizontal relations with the client should prevent the vertical power relationship from being turned into a closed iron cage.

Figure 4. A possible unit of analysis for examining power relations at work

Wolf-Michael Roth and his co-authors (2008) call for the inclusion of sensuous aspects of work into the unit of analysis. They name emotions, identity, and ethico-moral dimensions of action as salient sensuous aspects. Roth suggests that the sensuous aspects may be approached by focusing on actions together with their effects. This is basically the same insight that drives Sannino’s (2008) analysis of conflictual discourse.

Analyzing actions together with their social and material consequences is indeed a promising way to approach emotions and other sensuous aspects of activity empirically. But it is also important to ask: Why emotions? What is their role in activity? For Leont’ev (1978), emotions were above all signals of the subjective construction of object-related motives that are difficult to access and explicate consciously. To gain access to motives, one must proceed along a “round-about way,” by tracing emotionally marked experiences (Leont’ev, 1978, p. 125). In
other words, the study of action-level emotional experiences is an avenue to an understanding of activity-level motives. Mäkitalo (2005) took this route in his study of employees’ work-related emotions in nursing homes. The analysis of emotional experiences led to the identification of motives and different emotionally significant objects, which led to the identification of historically different but co-existing layers of the work activity.

Third generation activity theory expands the analysis both up and down, outward and inward. Moving up and outward, it tackles multiple interconnected activity systems with their partially shared and often fragmented objects. Moving down and inward, it tackles issues of subjectivity, experiencing, personal sense, emotion, embodiment, identity, and moral commitment. The two directions may seem incompatible. Indeed, there is a risk that activity theory is split into the study of activity systems, organizations and history on the one hand and subjects, actions and situations on the other hand. This is exactly the kind of split the founders of activity theory set out to overcome. To bridge and integrate the two directions, serious theoretical and empirical efforts are needed.

Coming from the study of written communication, David Russell (in press) suggests ‘genre as social action’ as a unit of analysis complementary to the unit of activity system. For Russell, genres are classifications of artifacts-plus-intentions. They are links between subjects, tools and objects. Genres provide relatively stable ways of seeing what acts are available and appropriate in a given situation.

I see genre and activity indeed as complementary concepts, much like Bakhtin’s (1982) concepts of social language and voice may be seen as complementary to the concept of activity. The concept of genre is very flexible and open-ended. This is both a strength and a weakness. Perhaps the most serious limitation has to do with the strong anchoring of genre to writing and written text. Activities are mediated by multiple modalities, from bodily movements and gestures to pictures, sounds, tools, and all kinds of signs. Written text is but one of the mediational modalities. It is not clear to what extent the concept of genre can be useful for analyses of activities in which multiple modalities work in concert and interpenetrate one another.

What is particularly interesting about genres as systems of typified written communication is their mobility and ability to cross organizational boundaries. Printed forms, records, genres of email and
other forms of documentation travel across activity systems and make trails that change the landscape. This is directly relevant for our attempts to understand current historical transformations in the organization of human activities.

The recent rise of new forms of Internet-based social production, or “commons-based peer production” (Benkler, 2006, p. 60; see also Shirky, 2008) prompts us to rethink the shape of activity systems. Third generation activity theory still treats activity systems as reasonably well-bounded, although interlocking and networked, structured units. What goes on between activity systems is processes, such as the flow of rules from management to workers depicted in Figure 4. Processes are commonly assumed to be relatively straightforward, stepwise movements from point A to point B.

In social production or peer production, the boundaries and structures of activity systems seem to fade away. Processes become simultaneous, multi-directional and often reciprocal. The density and criss-crossing of processes makes the distinction between process and structure somewhat obsolete. The movements of information create textures that are constantly changing but not arbitrary or momentary. The textures are made up of traces or trails which are both cognitive, ‘in the mind’, and material, ‘in the world’ (Cussins, 1992). Wikipedia is a good example in that every alteration of an entry is automatically stored and retrievable for anyone as a cumulative record of previous versions and alterations. So the constantly moving texture is also multi-layered and historically durable.

I have characterized these new forms of activity as ‘wildfire activities’ and ‘mycorrhizae activities’ in which interaction takes the shape of knotworking without a single stable center (Engeström, 2006, 2007a, 2008). Although greatly enhanced and accelerated by the web, I don’t think they are necessarily dependent on the Internet. Perhaps the new grassroots cooperatives spreading in Latin America, described by Klein (see above), are to some extent also examples of this kind of organizing.

If largely invisible, weakly bounded textures of criss-crossing trails become the foundaion of an activity, will the model of an activity system become obsolete as a unit of analysis? It seems clear that social production or peer production does not eliminate more bounded and vertically structured organizational units. Mycorrhizae are symbiotic forms which require trees and plants to survive and spread. Similarly, social production requires and generates bounded hubs of concentrated
coordination efforts. Thus, Wikipedia has the Wikimedia Foundation which collects funds, oversees the operation and occasionally institutes new rules and controls. The Wikimedia Foundation has a small paid staff working out of a main office in San Francisco. The web page of the foundation even displays a classic vertical organization chart. Activity system models are very appropriate for the analysis of such hubs. The challenge is to integrate such analytical tools with new concepts appropriate for the analysis of trails and mycorrhizae. Perhaps this implies a need for a fourth generation of activity theory.

MEDIATION AND DISCOURSE

Georg Rückriem (in press) argues that activity theory as it presently exists is captive of the historically passing medium of print and writing. For Rückriem, the whole idea of mediation of specific activities by specific tools and signs misses the point of the ongoing societal and cultural transformation engendered by digital media, especially by Web 2.0. Mediation is an issue of the historically leading or dominant media. The entire scope and character of human activities is determined by the dominant media.

Rückriem is right that in much of activity-theoretical literature, probably including much of my own work, print and writing are taken for granted as the dominant cultural media. Such tacit assumptions may indeed blind us to the consequences and potentials of digital media.

If Rückriem is right, it is media that determine the nature and possibilities of human activity. This means that the object of activity is of secondary importance.

Here is disagree with Rückriem. I see Rückriem’s insistence on the decisive role of media as a particular form of technological determinism. His argument ignores what media are used for – what ends and objects they serve. Consequently, it also ignores the internal contradictions of objects in capitalism. To me the most interesting issues of Web 2.0 have to do with the aggravation of contradictions between exchange value and use value, between private ownership and public good, between proprietary and freely accessible or open forms of knowledge and production. While this aggravation is greatly facilitated by Web 2.0, it is not simply a consequence of digital media. Forms of similar aggravation are seen in struggles over the production and
distribution of generic drugs, or indeed in the struggles over the uses of land and other natural resources in Latin America as reported by Klein.

Sweeping technological determinism leaves little room for human agency in concrete activities. Focusing on contradictory objects in specific activities calls for new forms of agency. When we take a closer look at the uses of digital media, much of the mythical omnipotence disappears. Thus, Shirky (2008, p. 136) characterizes Wikis (as well as other elements in Web 2.0) as “a hybrid of tool and community.” This characterization fits well in the classic analytical vocabulary of activity theory.

We do also need new concepts to make sense of Web 2.0. For example the notions of ‘open’ and ‘closed’ have great potential although they remain theoretically underdeveloped for the time being. Perhaps more importantly, digital media make very problematic the Vygotskian distinction between tool and sign.

The creation of new activity is a process of reflective re-mediation. A mediating concept or device can open up an entirely new question and lead to the formation of a new object and a new activity. This kind of re-mediation is radically different from goal-rational theories of change. The limitation of goal-rational models of creation and change is that they require that the investigator or interventionist defines the desired outcome of the change effort at the beginning. This leads to a paradox: How can you create something new if you know ahead of time what it is? Re-mediation involves a shift from the predefined or ‘given new’ goal to an unexpected or ‘created new’ object (Engeström, 1987).

When categories are imposed upon people they often become iron cages which reduce and rule out possibilities. Such closed ‘stabilization knowledge’ (Engeström, 2007b) is commonly the result of exclusively empirical generalizations taught in schools as authorized ‘correct knowledge’ (Davydov, 1990). On the other hand, existing social categories can also be turned into discursive tools that generate new, emancipatory meanings when blended with new contents and new categories. The President of Bolivia, Evo Morales, never made a secret of (a) being indigenous and (b) being a former coca farmer. These ordinarily very constraining categories, when blended with the category of President, were turned into a strength, in fact into symbols of entirely new possibilities and potentials. Such transitions from ‘stabilization knowledge’ to ‘possibility knowledge’ are at the core of zones of proximal development. The zone is never an empty space to be
simply filled with the new. It is inhabited by previous categories that need to be opened up, challenged and transformed.

In radical transformations aimed at the creation of qualitatively new patterns of activity, opening up and blending existing categories are not enough. What is needed is re-mediation by new theoretical concepts that serve as ‘germ cells’ for expanded horizons of possibilities. Davydov’s (1990) idea of theoretical generalization has nothing to do with scientism which regards ‘scientific concepts’ as superior compared to everyday concepts. Davydov carefully showed that science as taught in schools is in fact dominated by empirical generalizations. The roots of theoretical generalization are in our primordial attempts to change our conditions and to experiment with new solutions.

**EXPANSIVE LEARNING AND DEVELOPMENT**

Development is a burdened, yet necessary concept. As Rist (2006, p. 10) put it, "the principal defect of most pseudo-definitions of 'development' is that they are based upon the way in which one person (or set of persons) pictures the ideal conditions of human existence.” He proposes an alternative notion of development, not based on an ideal end state but on a realistic observation of what is being done in the name of development.

"'Development' consists of a set of practices, sometimes appearing to conflict one another, which require - for the reproduction of society - the general transformation and destruction of the natural environment and of social relations. Its aim is to increase the production of commodities (goods and services) geared, by way of exchange, to effective demand." (Rist, 2006, p. 13)

While Rist’s realism is a useful antidote to the taken-for-granted teleologies often present in theories of human development, it does not give us much in terms of understanding the destructive and constructive mechanisms of development. I will suggest a set of potential mechanisms that may stimulate further work in activity-theoretical studies of development. These mechanisms are (1) living movement, (2) breaking away, (3) double stimulation, (4) stabilization, and (5) boundary crossing.

(1) In the tradition of activity theory, a key metaphor for development is that of a zone. Often the zone of proximal development is interpreted as a vertical step which leads to a higher stage or level. I find it more useful to think of the a zone as a terrain of activity to be dwelled in and
explored, not just a stage to be achieved or even a space to be crossed. The zone is explored by moving in it. The movement may take various directions and patterns. In craft activity, the dominant pattern was from the periphery toward the center. In mass production, the dominant pattern is linear. At present, we see the emergence of patterns of pulsation, swarming, and multidirectional criss-crossing.

(2) The dwellers create trails and the intersecting trails gradually lead to an increased capability to move in the zone effectively, independently of the particular location or destination of the subjects. However, the zone is never an empty space to begin with. It has pre-existing dominant trails and boundaries made by others, often with heavy histories and power invested in them. More than that, the existing trails, landmarks and boundaries are inherently contradictory, possessing both exchange value and use value, being both controlled by proprietary interests and opening up possibilities of common good. When new dwellers enter the zone, they both adapt to the dominant trails and struggle to break away from them. The latter leads to critical conflicts and double binds. The troublesome trail of a student through as mass university is an example, aptly characterized as ”an obstacle course” by Sannino (2005, p. 188).

(3) 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 agency and will.

(4) New trails and intersections are marked, stabilized and made durable mainly in three ways, namely by means of critical conflicts, by means of authority, and by means of reification into artifacts and conceptualizations. Critical conflicts are often seen as merely situational problems. However, as therapy researchers such as Vasilyuk (1988) have shown, conflicts can become durable emotional blocks or sources of recurrent irritation that restrict and channel the actions of human beings for years. Formation and execution of authority is an obvious source of stability, yet it is an issue barely touched by activity theorists thus far (I return to it later in this chapter). Reification into artifacts and concepts, the ”ratchet effect” as Tomasello (1999) calls it, is the
most visible and palpable form of stabilization.

(5) Boundary crossing occurs because human beings are involved in multiple activities and have to move between them. A school student must move from home to school to peer culture and back home. Boundary crossing also happens between collective activity systems and organizations, in partnerships and mergers but also in espionage and hostile takeovers. Boundary crossing provides material for double stimulation. It requires negotiation and re-orchestration. It is the most obvious aspect of the horizontal or sideways dimension of development.

These five mechanisms partly overlap the conceptual framework of expansive learning (Engeström, 1987, 2001). Obviously breaking away is closely connected to facing and resolving contradictions in the different steps of a cycle of expansive learning. And stabilization is closely connected to the construction of a new model and new tools for the activity. But expansive learning as a stepwise process of ascending from the abstract to the concrete by means of specific learning actions is not reducible to the five mechanisms. I see it as the sixth and most important mechanism of development.

There is much research and theorizing in developmental psychology that is compatible with the idea of development as 'breaking away and opening up'. What is missing is sustained research programs that would integrate the psychological, institutional and societal aspects of development, not only observationally and retrospectively but also proactively and by means of interventions. To make it more concrete, the emergence of new forms of work and organizational knowledge creation are still domains that seem to have nothing to do with core issues of developmental psychology. Yet, the cultural teleology of development is largely forged in the spheres of work, technology and organizational strategy. Development happens – and should be studied - in the forging of the future in politically and affectively loaded everyday discursive actions, decisions and change efforts.

**AGENCY AND COMMUNITY**

Authority is foundational for the sustained existence of a community – yet, as Taylor (in press) points out, there is no in-depth treatment of authority in activity theory. I would approach authority from a
historical point of view. In my recent book *From Teams to Knots* (Engeström, 2008), I try to capture something of the historical evolution of authority by means of a condensed table (Table 1).

Coordination is not exactly the same as authority. However, the achievement of coordination is a central manifestation of authority. Thus, it may be useful to think of organizational authority in terms of the dominant mode of coordination, including its tools.

Table 1. A historical sketch for conceptualizing authority, agency, and community

<table>
<thead>
<tr>
<th>NATURE OF OBJECT</th>
<th>LOCUS OF AGENCY</th>
<th>DOMINANT MODE OF INTERACTION</th>
<th>COORDINATING MECHANISM</th>
<th>LEARNING MOVEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRAFT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal object</td>
<td>Individual actor</td>
<td>Coordination</td>
<td>Identification and subordination</td>
<td>Peripheral participation, gradual transition toward the center</td>
</tr>
<tr>
<td>SOCIAL PRODUCTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problematic object</td>
<td>Team</td>
<td>Cooperation</td>
<td>Process management</td>
<td>Focal involvement, linear and vertical improvement</td>
</tr>
<tr>
<td>SOCIAL PRODUCTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Runaway object</td>
<td>Knots in mycorrhizae</td>
<td>Reflective communication</td>
<td>Negotiation and peer review</td>
<td>Expansive swarming engagement, multi-directional pulsation</td>
</tr>
</tbody>
</table>

The fourth column of Table 1 sketches the typical coordinating mechanisms in very broadly conceived historical types of production. In craft-based organizations, when each individual practitioner is focused on his or her own object or fragment of the object, practitioners are commonly held together by externally imposed or tradition-based identification and subordination. In industrial organizations, teams emerged as units for cooperative solving of problems. Their efforts are typically coordinated by various forms of explicit process management. However, teams run into troubles and find their limits when faced with objects which require constant questioning and reconfiguration of the division of labor, rules, and boundaries of the team and the wider organization – in short, negotiation across horizontal and vertical boundaries of the given process.

Negotiation is a central coordinating mechanism of the distributed agency required in knotworking within social production. Negotiation is required when the very object of the activity is unstable, resists attempts at control and standardization, and requires rapid integration
of expertise from various locations and traditions. Negotiation is more than an instrumental search for a singular, isolated compromise decision. It is basically a construction of a negotiated order (Strauss, 1978) in which the participants can pursue their intersecting activities. As Firth (1995, p. 7) put it, “in quite implicit ways, negotiation activity implicates the discourse process itself, revolving around such things as acceptability of categories used to describe objects or concepts, and the veridicality of facts, reasons or assessments.” Putnam (1994, p. 339-340) takes a step further and points out that successful negotiations tend to transform the dispute, not just reach an instrumental end.

“By transforming a dispute, I refer to the extent that a conflict has experienced fundamental changes as a result of the negotiation. Fundamental changes might entail transforming the way individuals conceive of the other person, their relationship, the conflict dilemma, or the social-political situation. ...In the transformative approach, conflicts are no longer problems to be resolved; rather, they are opportunities to create a new social reality, a new negotiated order, a different definition of a relationship, or a transformed situation.”

Social production, such as the Open Source software movement or Wikipedia, is dependent on constant, publicly accessible critical commentary and peer review. When peer review becomes reciprocal, open and continuous, it actually coincides with Putnam’s notion of transformative negotiation.

Authority and agency are closely related. In agentic actions, we gain authority and become authors of our lives. This happens within historically changing patterns of activity and mediation. In Table 1, historical change in the locus of agency is described as shifts from the individual to the team and further to pulsating ‘knots in mycorrhizae’. This does not mean that the agency of an individual subject disappears. It means that the individual faces new challenges in his or her attempts to attain the position of an agentic subject. These challenges may be characterized by means of the notion of relational agency (Edwards, in press). It seems clear that individuals engaged in multi-agency collaboration aimed at the creation of a new activity need to nourish and manifest relational agency in order to achieve, as a collective, the expansive agency necessary for the accomplishment of radical transformations. Relational agency and expansive agency are complementary lenses, one focused on the individual, the other focused on the distributed collective.

The analysis of agency is still in its infancy. We need to link Il’enkov’s (1977) concept of contradiction with Leont’ev’s (1978) concepts of
need, object and motive, and these further with concrete manifestations of will and agentic action. In between, there is space for intermediate concepts such as conflict, envisioning, identification, responsibility, experiencing, and commitment.

One gains authority and agency by being recognized by a community and by receiving support from a community. The character of a community is to a significant extent determined by how open or closed is the shared object of the community. Table 1 implies that we are moving toward increasingly open, amoeba-like communities characterized by multi-directional swarming, weak boundaries and no single stable center. If this is the case, authority and agency may also grow in unexpected ways, as multiple simultaneous and interacting ‘minority influences’ (Moscovici, Mugny & van Avermaet, 2008) from the peripheries rather than as a single dominant majority influence from the center.

**INTERVENTIONS**

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://specialed.wordpress.com/2006/02/10/educational-research-the-gold-standard/](http://specialed.wordpress.com/2006/02/10/educational-research-the-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.

This predetermined and linear view of interventions is actually shared by much of the literature on design experiments. For example 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 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.

"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.)

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, 2008). 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).

Sociological intervention studies differ from educational ones in that there are usually no safe institutional walls to protect the intervention form the vagaries of the outside world. Perhaps this is why the linear view common to both ‘gold standard’ interventions and design experiments is much less easily adopted in sociology. A good case in point is the work of Norman Long.

“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 both ‘gold standard’ interventions and 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)

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.

“There is not, 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 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. The crucial differences are:

(1) 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) construct a novel solution or novel concept the contents of which are not known ahead of time to the researchers.

(2) In linear interventions, the subjects are expected to receive and implement 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.

(3) In linear interventions, the aim is to control all the variables and to achieve a standardized intervention module 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
can be used in other settings as tools in the design on locally appropriate new solutions.

Vygotsky himself described the method of double stimulation 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 as 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 you 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.

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. Institutional activity systems such as schools and workplaces are bombarded by interventions from all kinds of outside agents (e.g., consultants, administrators, customers, competitors, partners, politicians). And inside the activity system, practitioners and managers incessantly make their own interventions. Thus, taking the notion of intervention as a starting point is a way to remind us that we as researchers should not expect nicely linear results from our efforts.

Activity theory takes the subjects, the participants, the local practitioners, very seriously. But 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). This methodological principle is independently and
imaginatively developed in the French methodology of the Clinic of Activity presented by Yves Clot (in press).

In our Change Laboratory interventions (Engeström, 2007c), such a dialogical and longitudinal relationship forms the foundation for 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.

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. His view of the process of theoretical generalization may be summarized with the help of Figure 5. 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.

Figure 5. Summary of Davydov’s view of theoretical generalization (Davydov, 1982, p. 42)

As I have pointed out before Davydov’s theory is oriented at learning processes within the confines of a classroom where the curricular contents are determined ahead of time. This probably explains why it does not explicitly contain critical questioning of existing dominant practices and concepts. Similarly, the last actions of Davydov’s model do not clearly imply the construction of culturally novel material practices. In my theory of expansive learning, the beginning and the end of the process of ascending from the abstract to the concrete are conceptualized differently (Engeström, 1999c).

Significant change is not made by singular actors in singular situations but in the interlinking of multiple situations and actors accomplished by virtue of the durability and longevity of objects (Engeström, Puonti & Seppänen, 2003). This calls for a conscious expansion of attention
beyond the subjects, to include and center on the objects of work and discourse.

Expansive learning is above all stepwise expansion of the object. The potential for such expansion is best discovered by means of interventions which open up the zone of proximal development of the activity system. The study of expansive learning in complex settings requires a longitudinal intervention approach which may be crystallized in the form of three methodological rules (Engeström, Engeström & Kerosuo, 2003): (1) follow the objects of activity in their temporal and socio-spatial trajectories, (2) give the objects a voice by involving the clients or users in dialogues where the object is made visible, articulated and negotiated, (3) expand the objects by organizing intervention sessions and assignments where the producers and clients construct new shared models, concepts and tools to master their objects.

The Change Laboratory sessions are a purposeful blend of elements familiar from existing practices and new elements brought in by the researchers. They are designed to serve as microcosms where potentials of co-configuration and knotworking can be experienced and experimented with.

“A microcosm is a social testbench and a spearhead of the coming culturally more advanced form of the activity system. ...the microcosm is supposed to reach within itself and propagate outwards reflective communication while at the same time expanding and therefore eventually dissolving into the whole community of the activity.” (Engeström, 1987, p. 277-278)

In practice, the methodological rules sketched above mean that selected objects of activity in the research settings are first followed ethnographically. Critical incidents and examples from the ethnographic material are brought into a series of Change Laboratory sessions to stimulate analysis and negotiation between the participants. These formative interventions also allow us to observe the formation of the subjectivity of participants, not only as subjects of their central activity (e.g., work), but also as subjects of learning and design, and finally as subjects of self-change or experiencing. The Change Laboratory sessions themselves are videotaped for analysis and used as stimuli for reflection. The participants of the sessions engage in constructing shared models and tools to enhance their collaborative mastery of the object. The objects are again followed as the new tools and models are being implemented. The procedure allows for the collection of rich longitudinal data on the micro-interactions and
cognitive processes involved in expansive learning as the participants make visible their work, moving between actions and activity, between the past, the present, and the envisioned future.

Radical overall transformations of activity may be often be beyond the reach of research-based interventions. So perhaps we might use our energies in smaller and more accessible change efforts? This suggestion seems to run counter to my emphasis on global runaway objects. However, my tentative conclusion was: ‘We need intermediate runaway objects which are less spectacular and more inviting.’ This is indeed a task for activity theory: Bring together the big and the small, the impossible and the possible, the future-oriented activity-level vision and the here-and-now consequential action.

REFERENCES


