The question of whether human thinking can come to simulate computer intelligence—i.e., AI in reverse—is addressed in this paper. Examples are given of three computer tools which perform several functions that constitute an intellectual partnership between student and tool. Such functions include: (1) assuming part of the intellectual burden in experimentation; (2) presenting learners with novel alternatives; (3) freeing the individual from lower level operations; (4) displaying intermediate processes; (5) providing intelligent tutelage; and (6) providing models of information representation, processes, and strategies. Conditions necessary for internalization of artificial intelligence functions are described, i.e., comprehensibility, generalizability, novelty, utility, explicitness, and mindful abstraction. An experiment in which 74 Israeli seventh graders were divided into three groups and interacted with a version of the software incorporating metacognitive guidance and content-specific questions (experimental group), questions only, or neither is then reported. Results show that the experimental group devoted the most time to reading the texts and scored highest on a posttest; these findings support the hypothesis that interaction with a tool that models reading-related metacognition leads to improved reading of new texts in another non-computerized setting. Implications of these findings for a cognitive-developmental theory, a cultural-communicational theory, and educational theory and practice are examined. (28 references) (MES)
The issue

It can be said that language makes us human, literacy makes us civilized (Olson, 1986), and technology makes us powerful. Indeed, technology empowers us in a range of domains, ranging from the physical to the medical, from the perceptual to the communicational, and from the athletic to the artistic. But does technology, particularly computer technology, empower our intellectual capacities? Does computer technology make us better thinkers, better learners, better problem solvers?

To be more specific, consider computer tools, the nature of which I will discuss below, that incorporate a modicum of artificial intelligence: they "represent and use knowledge to perform complex reasoning tasks - tasks typically associated with "intelligent" human behavior" (Hayes-Roth & Thordyke, 1985, p. 231). Such artificial intelligence is often designed to simulate thought processes, knowledge representations and problem solving operations of experts. In other cases, the artificial intelligence involved is assumed to simulate the kind of intelligence that could help students better accomplish learning tasks and achieve better comprehension of the material. Such artificial intelligence is, of course, quite novel to the students exposed to it as it entails new ways of data representation, new kinds of strategies, even new kinds of mirrors of one's own solution traces (J. S. Brown, 1984). For the students interacting with such tools the intelligence exposed to is indeed "artificial".

The question I wish to raise, then, is not whether computer "intelligence" can simulate human thinking, but rather whether human thinking can come to simulate computer "intelligence"? This is what I call "artificial intelligence in reverse".

Before you dismiss this question as too outlandish, let me mention some relevant theoretical and empirical work that may support it. The theory that comes most readily to mind in this context is of course Vygotsky's. According to Vygotsky (1978) "every function in the child's cultural development appears twice, on two levels. First on the social, and later on the psychological level;
first between people as an interpsychological category and then inside the child, as an intrapsychological category. This applies equally to volunatry attention, to logical memory, and to the formation of concepts. All the higher functions originate as actual relations between individuals" (1978, p. 52).

Recent work by Brown and her associates (e.g., Brown & Ferrara, 1985; Brown & Palinscar, in press) with cooperative learning is very much based on this theory with quite dramatic results. The point of this work is that cooperative, externally provided guidance in the process of reading, when done within what Vygotsky has called the "zone of proximal development", is internalized and becomes part of the child's cognitive repertoire.

However, it is not only social interaction that figures in a child's zone of proximal development. Also auxiliary tools and symbol systems which initially serve as means to affect objects of activity or represent them communicationally are internalized and come to serve as cognitive signs for self-guidance (Vygotsky, 1978). Indeed, in my own past work (e.g., Salomon, 1979), inspired both by Vygotsky and by Bandura's observational learning theory, I have found that students can internalize certain symbolic forms of the media and come to use them as transferrable internal modes of information representation and manipulation. Could not the same be applicable to the kind of artificial intelligence one interacts with when using computer tools? Could not such tools serve as "more capable peers" (Vygotsky, 1978) the functions of which are then internalized?

Computer tools

To be sure, computer technology offers learning activities that amplify existing practices and serve time-honored instructional goals. Such usages of the computer may expedite the acquisition of that knowledge students were always required to learn. But there are other opportunities that computer technology offers which, as I will try to show, might have qualitatively different effects on students' intellectual performance, even on their intellectual capabilities. I am speaking of a class of computer-related activities that are based on computer tools. Here, unlike other cases, and perhaps for the first time in instructional history, students are given the opportunity to enter into an intellectual partnership with relatively intelligent tools whose capacities can transcend the limitations of human cognition (J. S. Brown, 1984; Pea, 1985; Salomon, 1988).
What are these tools and what constitutes this "intellectual partnership"? Consider by way of example three tools. The Learning Tool (Kozma & Van Roekel, 1986) allows learners to create spatial maps of concepts, events or dates, and to interrelate them by a variety of lines that represent causal, part-whole, temporal or correlational links. Each concept in the "map" can constitute a "map" of its own in a hierarchical structure, affording top-down and bottom-up review of the material. Another example is Ecolife (Mintz, 1987), an instructional simulator by means of which children can manipulate a variety of ecological variables, thereby creating dynamic "environments" the development of which they can then explore through pictorial, graphic and numerical displays.

A third example is Stella (Doyle, 1985), designed for the construction of dynamic, complex, mathematically based models. The student selects a number of, say, economic, ecological, dietary, or transportation variables and represents them in spatial form. Then, the student decides on the kinds of mathematical relations among the variables and the initial values for the critical one. Once built into the model, the designed "environment" starts "operating", producing dynamic graphic and numerical representations of the "environment".

Now, such tools, although still limited, accomplish a number of functions. First, they assume part of the intellectual burden that is involved in the process of generating and testing hypotheses, writing, designing, planning, conducting experiments, and the like. The part that such tools carry out for the learner are usually the tedious, lower level ones that often exceed one's cognitive capacities: computation, rearrangement of display, drawing. In accomplishing this, tools can be said to serve as extensions of our cognitive functions.

Second, tools encounter learners with novel choice points or alternatives, the resolution of which requires the mobilization and acquisition of knowledge (Olson, 1976). For example, the Learning Tool requires one to consider not so much the list of constructs or events to be mapped, but the way they interrelate.

Third, tools accomplish an enabling function; freeing the individual from lower level operations, they enable the mindful individual to test new possibilities and examine their consequences. How would the economic results
look once I change the configuration of market variables in the Stella model which I construct?

Fourth, tools can display intermediate states and processes en route to the final solution or design. They serve in this respect as trace systems whereby "the intermediate products of mind are externalized" (Pea, 1987, p. 138).

Fifth, they can (although tools for the professional rarely do it) provide intelligent tutelage, guidance and probing - by raising questions, signalling errors, suggesting moves, or providing externalized metacognitive-like guides.

Last, and perhaps of greatest importance, tools often provide explicit models of information representation (like the map in the Learning Tool), of processes (as in the case of the developing graph in Stella), or of strategies (as in the case of the Reading Aid).

**Intellectual partnership**

Working with tools that accomplish at least some of these functions entails three important ingredients: it entails a *division of labor* which is *complementary* and through which the operations involved and the products evolved become *interdependent*. These are the same three qualities one finds in human partnership. But this partnership is of a special kind: it is an *intellectual* partnership insasmuch as the complementary division of labor, and the interdependence of processes and intermediate products involved, concern intelligent activities with symbolic entities, activities that if carried out by an individual without that partnership would strain, even exceed, the limits of human cognition (Pea, 1985).

What about this partnership? It can be expected that this partnership has the potential of profoundly upgrading students' performance (Salomon, 1987): think of the lab experiments students can perform with the Lab Patner (Linn, Layman, & Nachmias, in press), the kinds of hypotheses in physics that students can generate and explore with a *Newtonian Computer Microworld* (White, 1984), or the kinds of relationships they can discover when mapping a field of study by means of the Learning Tool. In this respect, one could say that the partnership-like system of learner and computer is now more intelligent (Pea, 1985, 1987). Intelligence, in this case, is not seen any more as a quality of an individual's mind alone but rather as "a product of the relationships between mental structures and the tools of intellect provided by the culture"
(Bruner, 1966, cited by Pea, 1985, p. 168). Indeed, as Pea (1987) argues, "Just as the human body is no longer the major tool for physical labor, and just as a carpenter need not use only hand tools, so will mental functioning no longer be the sole province of the human mind" (p. 144). So, does computer technology make us better thinkers or learners? For a first cut, the answer could easily be positive, given the intellectual partnership just described.

**Cognitive residue**

However, for a second cut, the question I had in mind was not whether so-called cognitive tools upgrade intellectual performance during the partnership. Rather, my question pertained to the possibility that this partnership leaves some cognitive residue - affecting cognitions that can be used without the scaffolding of the external, intelligent tool. Paraphrasing Vygotsky, we might say that the partnership creates a zone of proximal performance. This zone defines the difference between what learners can do on their own and how well they can perform when accompanied by an intelligent tool. This is not necessarily a zone of proximal development. For the latter "defines those functions that have not yet matured but are in the process of maturation, functions that will mature tomorrow" (Vygotsky, 1978, p. 86, emphasis added). Once matured through interaction with external agents, these functions should become "part of a child's independent developmental achievement" (Vygotsky, 1978, p. 90). That is, one should be able to observe them outside the guidance offered in the partnership with the tool.

I need to clarify what I mean by the child's operating without a tool's guidance. Computer tools, I have argued, accomplish a number of functions. One is the enabling function; others are the guidance and modeling functions. The latter, not the former, are of course the candidates for internalization. The extent to which such internalization has actually taken place could, however, be determined only when the tool's guidance and modeling, but not its enabling features, have been removed. For it is in the presence of the latter that a child could employ the internalized strategies or modes of representation to their fullest.

It follows from a Vygotskian perspective that improved performance with a tool is a necessary step in achieving a more lasting cognitive change as a consequence of interacting with it. But it may not be a sufficient condition. Thus, for example, using a word processor that provides externalized
metacognitive guidance during the writing process may significantly improve the quality of the essays written, but this does not necessarily equip the writer with better writing strategies, manifested later when using a regular word processor, let alone a pencil and paper. The employment of a tool needs to facilitate the internalization of certain artificial intelligence functions such that they become transferrable cognitive functions.

What would constitute evidence for the internalization of a tool’s strategies or modes of representation such that it serves as a cognitively reconstructed tool? First, one would want to see evidence that the external tool-based elements have been internalized, possibly reflected in the child’s new mastery of functionally or even structurally similar cognitive functions (Diaz, 1986; see discussion of the similarity issue by Wertsch & Stone, 1985). Second, to be considered as a cognitive tool, not just a highly domain-specific bit of knowledge, one would want to see the internalized function transfer to some other instances. Obviously, not every tool and not every encounter with a tool, intelligent as it might be, is likely to yield internalization of artificial intelligence that becomes a transferrable cognitive tool. A number of conditions need to be met to afford such processes. To these I turn next.

Some conditions for tool internalization

I have started out with the general hypothesis that when certain conditions are met, the "intelligent" strategies and modes of representation modeled by a computer tool can be acquired by a child to become cognitive. How would this take place? I have argued in the preceding section that improved intellectual performance achieved in cooperation with an external tool may be a necessary condition but that it does not suffice as it need not be accompanied by internalization, that is - be mentally reconstructed to create appropriate internal processes (Wertsch, 1985; Wertsch & Stone, 1985) applicable to novel instances.

I wish to suggest five conditions that need to be met for the operations of any computer tool to be internalized. First, for computer-based strategies and modes of representation to become cognitive they must be such that they could potentially be carried out in one’s mind. Thus, for example, children might be expected to internalize explicitly presented metacognitive guidance, but they would not be expected to do the same with long computations. The latter are likely to exceed children’s cognitive capacities.
The second condition suggests that the nature of the encountered strategies, the computer "intelligence" that is candidate for internalization, should be of a generalizable nature, that is - potentially applicable to new instances. For example, the conceptual mapping afforded by the Learning Tool can be applied to a wide range of topics and instances, not so the formulae of a spreadsheet, which cannot.

Third, the "intelligence" encountered through the partnership with the tool should serve a novel and useful cognitive function. Children who can already perform relevant tasks, efficiently employing their own strategies, are less likely to internalize the tool's strategies than children for whom the strategies carry novelty and utility (Salomon, 1974). This is where computer-based intelligent tools provide a unique opportunity. They offer strategies and modes of representation hardly ever encountered elsewhere.

The fourth condition suggests that the tool's "intelligent" operations be explicit, such that the child can actually witness, trace and reproduce them. Strategies employed by a tool "in hiding", as is the case with the mathematical operations of, say, a spreadsheet, cannot be encountered and certainly not mentally reconstructed. Nor would one expect a child to internalize strategies from a tool that employs none. Typical word processors may afford certain writing strategies but they do not "process" the written material in any observable strategic way. It is with respect to this condition that I have argued that a tool can lead to better performance and yet fail to serve as a source for strategy internalization: It carries out the crucial operations away from child, not allowing the child to interact with the very operations or strategies to be internalized.

The fifth condition does not pertain to the tool but to the way it is interacted with. Based on the distinction between automatic and controlled processes (e.g., Fisk & Schneider, 1984), it has been argued (Perkins & Salomon, 1987) that the acquisition and transfer of a strategy can take either one of two distinct routes, or a combination thereof. The first route is by repeated practice with a variety of problems which leads to proficient mastery of the strategy and to its automatic employment. Being already automatic, this implies that the new situations to which the strategy becomes transferred bear great perceptual similarity to the
situations in which it was acquired and practiced (think of how automatically
driving a car transfers to trucks but not to semi-trailers). We have called this
way to transfer the low road (Perkins & Salomon, 1987).

While the low road to transfer can account for the automatic application
of well mastered mental skills, tacit knowledge, attitudes, culturally acquired
habits, and the like, it cannot account for the cases of transfer of disciplined
knowledge acquired in school when practice to near automaticity is not
provided (e.g., Bransford, 1987). Such transfer appears to take the high road
which entails the volitional employment of controlled processes. Specifically,
it entails processes of mindful abstraction, that is, deliberate, effortful and
metacognitively guided decontextualization of a principle, main idea, strategy,
concept or rule (Salomon & Globerson, 1987). Mindful abstraction, the essence
of the high road to transfer, comes close to what Vygotsky has labeled
"intellectualization". It is the decontextualized principle or strategy that
becomes applied to new and perceptually dissimilar instances. Think, for
example, of how a chess player might mindfully abstract from the game the
strategy of controlling the center and apply it to, say, a game of basketball.
Just repeatedly practicing chess is unlikely to yield such a decontextualized
strategy and certainly not lead to its application to anything outside of chess.

Interaction with a computer tool can take either the low or the high
road. One can use a tool to solve problems, create models, write, read, or design
"environments", without being mindful of the strategies employed or
manifested by it. Such a user may internalize the explicit strategies but this
requires much and prolonged practice, not often provided to tool-using
children. This user is not likely to employ the internalized strategy unless it
has become fully automatized. Even then, it might not transfer very far. To the
extent that the modes of operation of so-called "defining thechnologies"
(Bolter, 1984) have gradually come to affect the human mind over the decades
and millenea, it would have been by means of this low road of unnoticed,
practice intensive internalization. Alternatively, the tool-using child may be
(or made to be) mindful of the ways in which the tool operates on the data,
allowing him or her to mindfully abstract a strategy from these encounters
and internalize it, that is - mindfully reconstruct it in his or her mind. To the
extent that this condition is met, the likelihood of transferring the
internalized strategy to novel situations and tasks would be increased.
An experimental illustration: The Reading Aid

Given that all these conditions are met, can a tool's logic, its strategies, or its modes of data organization, be internalized and come to serve as part of a child's cognitive repertoire? Salomon, Globerson, & Guterman at the Tel-Aviv University have designed a computer tool, the Reading Aid, to test the hypotheses presented here. In what follows I will briefly describe the tool and the study.

The Reading Aid is a prototypical computer program which accomplishes two functions. First, it presents selected strategies relevant to the reading process - reading titles and making predictions on their basis, the generation of images of the text, the detection of key sentences, and the construction of subsummaries. Second, it provides explicit metacognitive guidance during the process of reading texts. These externalized metacognitions are akin to the ones known to accompany the reading processes of proficient readers (e.g., Baker & Brown, 1984), modelling the kind of metacognitive guidance children would be expected to subsequently use on their own. Examples of the metacognitively guiding questions used by the tool are "what kind of image have I created from the text?", "what thoughts occur to me on the basis of the title?", "what do I understand from the text so far?", and the like. The Reading Aid was designed for sixth and seventh graders.

Seventy four Israeli seventh graders interacted with the Reading Aid during three two-hour sessions. In each session, one strategy was presented followed by the reading of three or four computer-displayed texts, eleven texts altogether. These texts were chosen from among a large number of texts on the basis of their interest and readability. During the reading of a text metacognitive question would appear in a "window", one question at a time. The children were encouraged to read the questions carefully and to mindfully answer them to themselves. They then "erased" the question by hitting a control key (the computer recorded the time each question was displayed on the screen).

We contrasted the Reading Aid with two other versions. One version entailed no externalized metacognitive guidance but presented content-specific test-like questions at the end of each of the 11 texts. The children were
encouraged to mindfully answer these questions for themselves. The other, control version, consisted of the 11 texts to be read with no guidance and no content-specific questions. The children were randomly assigned to each of the versions, 24 to 25 children in a group. The reading sessions were held in mixed groups of about 12 children, each child working alone.

A number of pretest, process and posttest measures were taken (we report here only the ones most pertinent to the present discussion). These included a pretest of reading comprehension (a standardized test used in Israeli schools); average reading time of the texts; mindfulness during the reading (measured by means of self reports of effort expenditure and "how much I was thinking about the way I was reading", administered at the end of each session); and three kinds of posttest: children's ability to offer (metacognitive) guidance they would "give to a friend who does not read very well"; a delayed (10 days) reading comprehension test (another form of the pretest); and a delayed (one month) transfer task - essay writing.

Results were consistent and clear. There were no initial differences in reading comprehension between the three groups, however large differences emerged on the process measures. The experimental (Reading Aid) group devoted significantly more time to the reading of the texts than either one of the two other groups, F(2,71) = 16.80, p < .001, accounting for 32% of the reading time variance. The experimental group also reported being more mindful during the process of reading, with the other groups trailing behind it, F(2,71) = 2.86, p < .06, accounting for 12% of the variance. Combined, these findings suggest that, on the average, the experimental group was not just passively exposed to the guidance but mindfully engaged in answering the metacognitive questions.

A similar difference was found on the measure of metacognitive reconstruction, a measure we expected to reveal the extent to which the externalized guidance has been internalized: F(2,71) = 30.31, p < .001, accounting for 46% of the variance (Figure 1).

This in itself is not surprising as only the experimental group was exposed to the explicitly presented metacognitions. However, the other
findings showed that this difference between the groups was anything but trivial. When reading posttest scores were compared (partialling out pretest scores), we found again that the experimental group attained the highest scores, F(2,71) = 4.95, p < .05, accounting for 17% of the posttest reading variance (Figure 2).

But was this improvement related to the internalized metacognitions? When posttest reading comprehension scores were compared again with reconstruction of metacognitions partialled out, the difference between the groups totally disappeared. The difference in metacognitive reconstruction was not trivial at all; ability to reconstruct reading-related metacognitions accounted for all the improvement in reading comprehension.

These findings supported our hypothesis that interaction with a tool that explicitly models reading-related metacognitions leads to improved reading of new texts in another, non-computerized setting. Neither practice in reading nor the provision of content-specific review questions led to such an improvement. However, these findings do not show much far transfer. Analysis of the essays written one month later revealed exactly that. The essays were independently graded along a number of criteria by trained graduate students blind to the study's hypotheses. Comparisons of the essays' overall quality revealed that the experimental group wrote significantly better essays than the other groups, F(2,71) = 3.92, p < .05, accounting for 9% of the variance (Figure 3). As before, partialling out scores of metacognitive reconstruction nullified the difference in writing quality. Relatively far transfer has thus been shown.

Discussion

This study demonstrates quite clearly that a computer tool, even when particularly "intelligent", can become cognitive and that cognitive ability can result from the mindful internalization of its explicit strategies.
The study shows that a computer tool even when not (yet) very "intelligent", can serve as a "more capable peer", offering guidance in a child's zone of proximal development, guidance that can be internalized and used as a relatively generalized cognitive strategy.

The possibility that artificial intelligence can become cognitive may raise a serious ethical question. Are we not trying to infuse young minds with truly artificial, machine-like logic which might supplant their own natural logic? Perkins (1987) has dealt with this issue when writing about the cultivation of particular "thinking frames". These, like the artificial intelligence dealt with here, can be claimed to be alien to natural thought processes, cultural artifices, if you will. The answer given by Perkins (1987) and, unrelatedly by myself (Salomon, 1988) was that human thought has always been infused with so-called artificial processes and strategies that run contrary to the human grain: thinking about and redefining a problem rather than aiming right away at its solution; callenging basic assumptions; thinking about both sides of a case; and the like. And, as Olson (1986), points out, also the ability to distinguish what is said in a text from what is meant and attributed to it, was initially forced upon the Western mind by the development of literacy and later by the spreading of print. Natural stupidity was always, throughout history, enriched by more powerful "artificial" intelligence which gradually became part of more natural intellectual functioning.

But there is a difference between the thinking artifices Perkins (1987) alludes to and the ones one is exposed to when working with a computer tool. The former are acquired and used by the individual when on his or her own; the latter are acquired during an intellectual partnership in which much can be left to the tool to be carried out. Hence, a word of caution is in place. Computer tools, I have argued, can become cognitive as a result of the partnership that evolves between them and learners, a partnership in which abilities, so to speak, are pooled. However, it might also be the case that a highly intelligent tool makes less-than-mindful users become dependent on it, not taking full advantage of the tool's enabling possibilities, thus relinquishing the employment of their own abilities. Why should one emulate the logic of, say, an expert system when this system is so efficient, clever and easy to use? The learner's mindfulness concerning the way the tool operates, the logic it displays and the opportunities it affords, appears to be a factor of prime importance here (Perkins, 1987).
With this caution in mind, what are the implications that can be drawn from the arguments and the findings I have presented? First, there are implications for a cognitive-developmental theory, providing evidence for a Vygotskian approach to development and to the role that instruction plays in it. One could now add to that theory the conditions under which also intelligent computer technology can affect thought processes. But this extension of a Vygotskian theory also raises new questions. Are computer's effects and its mechanisms identical to the ones observed when humans interact, say, in cooperative learning (e.g., Palinscar, Brwon & Martin, 1987)? May be not: unlike cooperative learning which is temporary, lasting only until a set of skills has been acquired, one's partnership with a computer tool is continuous. What are the cognitive-developmental implications of such an ongoing partnership?

Second, there are implications for a cultural-communicational theory. These implications shed light on the way communicational media and tools affect minds, not only on a socio-cultural scale, as studied for example by Havelock (1982) or Ong (1982), but on a much shorter time scale, as discussed by Olson (1986). These effects pertain to the way individuals' intelligence develops. The ways in which cultural artifacts affect minds on a societal scale - through shared metaphors (Bolter, 1984) or shared cultural representations (Sperber, 1984) - may not be the same ways in which they affect the individual mind.

Third, there are implications for educational theory and practice, concerning ways in which computers can be used in schools. Computer tools, and the intellectual partnership they afford, may not only upgrade performance and facilitate learning, but they can also become relatively novel cognitive tools. Should this then shift the balance in favor of the development and use of intelligent tools at the expense of, say, drill and practice? Should the possibilities described here come to guide tool design? Should we aim more at the creation of what I have called zones of proximal performance, facilitating intellectual partnerships to upgrade performance, or should we see in them means for the creation of zones of proximal development to facilitate tool internalization? Tools, let us remember, do not just ease the tasks we would have undertaken anyway; they redefine the nature of the tasks, even redefine our relationship to the world we encounter and to our own cognitions. Tool and mind affect each other reciprocally.
References


Note

The research reported here was supported by a grant from the Spencer Foundation. This paper is based on two other papers by Salomon (1988) and by Salomon & Globerson (submitted for publication to LCHC).
Figure 1:
Reconstruction of Metacognitions – Posttest

- Metacog. Group
- Content-Questio
- Control
Figure 2: Pre- and posttest reading scores
Figure 3: Writing posttest scores

- Metacog. Group
- Quest. Group
- Control Group