MODELING THE PERCEPTUAL COMPONENT OF CONCEPTUAL LEARNING—A COORDINATION PERSPECTIVE

William J. Clancey
Institute for Human and Machine Cognition and
NASA-Ames Research Center

Appeared in P. Gärdenfors and P. Johansson (Eds.), *Cognition, Education and Communication Technology*, Lawrence Erlbaum Associates, 2005, pp. 109-146. Presented at symposium in Stockholm, April 2000. This is a preprint; please consult the publication version when citing.

> Visual forms are not discursive. They do not present their constituents successively, but simultaneously, so the relations determining a visual structure are grasped in one act of vision. Their complexity, consequently, is not limited, as the complexity of discourse is limited, by what the mind can retain from the beginning of an apperceptive act to the end of it. Of course such a restriction on discourse sets bounds to the complexity of speakable ideas. An idea that contains too many minute yet closely related parts, too many relations within relations, cannot be "projected" into discursive form; it is too subtle for speech. A language-bound theory of mind, therefore, rules it out of the domain of understanding and the sphere of knowledge.
>
> Susanne Langer, 1942, p. 86

Although a picture may be worth a thousand words, modeling diagrams as propositions and modeling visual processing as search through a database of verbal descriptions obscures what is problematic for the learner. Cognitive modeling of language learning and geometry has obscured the learner's problem of knowing where to look—what spaces, markings, and orientations constitute the objects of interest? Today we are launching into widespread use of multimedia instructional technology, without an adequate theory to relate perceptual processes to conceptual learning. Does this matter? In this article, I review the symbolic approach to modeling perceptual processing and show its limitations for explaining difficulties children encounter in interpreting a graphic display. I present an alternative analysis by which perceptual categorization is coupled to behavior sequences, where gesturing and emotional changes are essential for resolving impasses and breaking out of loops. I conclude by asking what kind of cognitive theory we need to exploit communication technology. Have we been correct to assume that pedagogy must be grounded in an accurate psychological model of knowledge, memory, and learning?

THE ZBIE MODEL OF LANGUAGE LEARNING

In a pioneering computer model, Siklossy developed a model of language learning that Vera and Simon (1993) believed refuted the claim that knowledge is not stored as descriptions in the human brain:

> Clancey [1993b] says, "Regularities develop [in behavior] but without requiring us to represent them as rules or graphic networks or pictures. The obvious example is of a child learning to speak before being taught an abstract grammar" (p. 103).
>
> Now the best example we know of a theory of how a child learns to speak is the program, ZBIE, written by Siklossy (1972) which does just that. It learns language by seeing sentences in juxtaposition with the scenes they denote, and gradually acquires both vocabulary and grammar,

> together with the ability to produce sentences never before experienced, when presented with new scenes. The grammar (stored in the simulated child's memory) is not in the form of rules that the child is aware of and can state; it is in the form of active procedures that are gradually built up through experience as an integral part of a changing program—all done with a purely symbolic representation. An empirical demonstration of a phenomenon provides a convincing refutation to a "proof" of impossibility. Can any existing SA theory perform this learning task? (p. 128).

Let us look at ZBIE and see whether it fits Vera and Simon's claims. To begin, Siklossy tells us that the inspiration of ZBIE's design is I. A. Richards language-through-pictures series:

> ...pictures are associated with sentences in an NL [Natural Language] to be learned. The pictures are to act as a general representation that has uniform meaning for all human beings (*English through pictures*, Book I is prefaced in 41 languages). The student is supposed to use the pictures as clues to the meanings of the sentences and, by successive comparisons of the sentences, to infer the vocabulary and grammar of the NL studied.
>
> The student's own mother tongue is bypassed, thereby avoiding problems of translation from one language into another; instead the student learns to translate situations directly from "reality" into a new NL.

Siklossy tells us that the idea of language-through-pictures is to associate meanings with pictures, rather than descriptions in another language. In this way, "the student's own mother tongue is bypassed." Instead of translating between languages, "the student learns to translate situations directly from 'reality' into a new NL."

This theory has some merit, though the learning process is not "translation." If there were sufficient context, such as in a cartoon strip relying on a common cultural background, a student could to some extent understand the meaning of the pictures and relate this to the words of the new language. Indeed, language learning must involve some aspect of coordinating non-linguistic conceptualizations (images, gestures, interpersonal relationships, etc.) with linguistic statements.

Amazingly, Siklossy notes parenthetically that language-through-pictures learning *does not work in his experience* or that of other people he knows:

> (As an aside, the author tried to learn Hebrew, absolutely unknown to him beforehand, from Hebrew Through Pictures. He had the advantage of having previously read several other known languages; nevertheless he had great difficulty in determining the meanings of the pictures or the clues to be derived from them, and finally abandoned the endeavor. Several other persons reported identical difficulties.) (Siklossy, 1972, p. 289-290)

If the pictures are insufficient, how could a computer implementation of this approach work? Simply use as computer input descriptions of the pictures, so the process is indeed translation and no visual perception is required. This model is ZBIE, though Siklossy glosses the difference from the original problem:

> The philosophies between ZBIE and I.A. Richards' booklets are similar. ZBIE uses a functional language (abbreviated FL) to represent situations; FL has the same function in ZBIE as the pictures have in Richards. By successive comparisons of situations, as represented in FL and as expressed in an NL, respectively, ZBIE tries to express other situations represented in FL and, failing that, to use its previous knowledge to learn how to express the other situations. The learning sequence presented to ZBIE is taken from Russian Through Pictures with slight modifications. (p. 290)

Table 1 illustrates how ZBIE represents a situation in FL, corresponding to the NL expression to the person is supposed to be comprehending.

Table 1. Formal Language (FL) representation for corresponding Natural Language (NL) statement describing a picture.

Language	Expression
Formal (FL)	(be (on hat table))
Natural (NL)	The hat is on the table.

Siklossy tells us that "FL has the same function in ZBIE as the pictures have in Richards." That is, the linguistic statement "(be (on hat table))" serves the same function of providing a "reality" to be related to words as a picture in Richards' booklets.

Here Siklossy, like Vera and Simon, makes no distinction between a description of a picture and the picture itself:

> It [ZBIE] learns language by seeing sentences in juxtaposition with the scenes they denote, and gradually acquires both vocabulary and grammar, together with the ability to produce sentences never before experienced, when presented with new scenes.

But ZBIE does not *see* scenes at all! ZBIE relates linguistic statements in FL to linguistic statements in NL. We have therefore gone from an assumption in Siklossy's paper that FL statements serve the same function as pictures, to a retelling in Vera and Simon's paper that ZBIE sees pictures themselves.

The distinction is crucial. The ZBIE model of language learning, which Vera and Simon call "the best example we know of a theory of how a child learns to speak," operates by mapping statements to one another. For this to be a model of language learning, as Vera and Simon claim, it is first necessary to learn (or have inborn) the functional language, FL, and to have a way of mapping perceptual categorizations to the words of FL.

Vera and Simon emphasize that ZBIE's manipulation of translation patterns corresponds to subconscious processes in a human being:

> The grammar (stored in the simulated child's memory) is not in the form of rules that the child is aware of and can state; it is in the form of active procedures that are gradually built up through experience as an integral part of a changing program—all done with a purely symbolic representation.

Although there is no distinction in the model between FL and NL statements and translation rules—all are statements—the interpretation of the model, according to Vera and Simon, is that the translation rules are not accessible to the person.

Although the idea of designing programs so they can "introspect" to read internal models and grammars has been the dominant approach for machine learning and expert systems design, Vera and Simon are correct that a scientist need not design a program in this way:

> This example also helps clarify a source of Clancey's confusion about conflation. He appears to be under the impression that any "rule" that appears in the computer memory of a simulation program (i.e., a production in the Category 4a symbol structures) must be accessible to the simulated person, and verbalizable as a Category 3 structure. Of course, this is false.

Granting Vera and Simon's point, the fact remains that there is nothing in Siklossy's theory of memory, perception, learning, or reasoning that distinguishes ZBIE's translation patterns from the FL and NL input: All three are statements in a language and all are manipulable by the program to produce new statements. That is, just as ZBIE can generate new NL statements and store them in memory, it can generate new translation patterns and store them in memory. Hence, the theory of comprehension (understanding FL or NL statements) is purely a process of mapping between and assembling statements. Nothing in the nature of the statements, their storage in memory, or how they are used prevents their accessibility.

Setting aside the issue that plagues all exclusively descriptive (symbolic) theories of cognition—how are the initial translation patterns learned—we must make sense of Siklossy's experience in attempting to learn Hebrew from Richards' booklet. How can relating pictures to NL (which Siklossy abandoned as too difficult for a person to do) be *equivalent* to relating FL to NL (which Vera and Simon claim is the best example of how a child learns to speak)? Siklossy's experience suggests that the process of language learning in human is surely not what ZBIE is doing. Rather ZBIE succeeds at its task because it is mapping between descriptions. Siklossy failed at learning Hebrew from pictures precisely because he was operating on *pictures*, not *descriptions of what the pictures represent*.

Indeed, to understand the relation of ZBIE to learning language-through-pictures, we must analyze the difference between Siklossy's and ZBIE's tasks in language learning more carefully.[1] Siklossy's task in using Richards' booklet involves attempting to determine what the pictures mean:

[PICTURE] <—> interpretation of picture <—> NL string

ZBIE's task involves mapping three *linguistic expressions* to each other in a syntactic parsing process:

FL string <—> formal translation rule <—> NL string

Siklossy's difficulty arises because the meaning of the picture is open to many descriptions. ZBIE's ease derives from being given a description to work with, "(be (on hat table))." Siklossy must create a suitable description by finding a way of viewing and conceiving of the picture—which are not exclusively linguistic processes. ZBIE is given the suitable description, indeed already in the vocabulary to be learned!

Siklossy acknowledges these limitations in his model. Furthermore, he points out that the FL representations (as well as the pictures in Richards' books) capture "an Indo-European's 'vision of the world.'" (p. 322), and as such are not a neutral, universal input by which a student who speaks any language at all could understand the target language.

But according to the symbolic view, the cultural variation of interpretation is merely a matter of different background knowledge. Which is to say, the initial patterns in ZBIE affect what the program can do. We are back to the same dispute: The symbolic view says that background knowledge consists of more descriptions. An alternate view is that the background knowledge consists of *ways* of seeing, speaking, and conceptually coordinating activity—not just descriptions of these (Clancey, 1997b, 1999).

How can we step outside of this argument? How can we show that human comprehension of text or pictures depends on a conceptualization which itself cannot be reduced to descriptions? Rather than starting with descriptions of arrangements—forms that a teacher claims are significant—we examine how students create their own significant forms. We study the process by which people perceptually and physically segment space to create forms that they claim are meaningful. Fundamentally, we examine how this process of "viewing as" and interpreting is inseparable in human experience, so seeing something meaningful and conceiving what it means

[1] Ironically, having inaccurately described ZBIE, Vera and Simon fail to see how my conclusions depend on careful analysis of existing computer programs. Instead they say: "These are examples of another mode of fallacious argument in which Clancey indulges. In this case, his arguments are simply based on misconception about the actual construction and operation of symbolic systems." (Vera and Simon, 1993, pp. 128-129).

occurs together, and is only subsequently followed by a coherent linguistic statement by which the meaning is represented.

In the following example, we turn around Vera and Simon's challenge: Can any existing computer program do what these children accomplish?

HUMAN LEARNING: A GREEN GLOBS EXPERIMENT

In this example, two students are working on an exercise intended to teach them about the properties of linear equations. The students (Pam and Susanna) are following and completing a worksheet that directs their use of a graphing program (called Green Globs[2]). Along the way, the students get confused about what a straight line is and miss the intended point of the lesson.

Asked to predict how the line for Y = 5X + 1 will appear, P&S first write "that the equation is going to get thicker." But after seeing the screen, they modify the answer—"not" is inserted above and before "going," and "thicker" is smudged out above "straight"}

What do you think will happen if you type in **Y = 5X +1**?
That the equation is not going to get ~~thicker~~ straight.
Sketch your prediction on this empty graph and then try it on the computer.

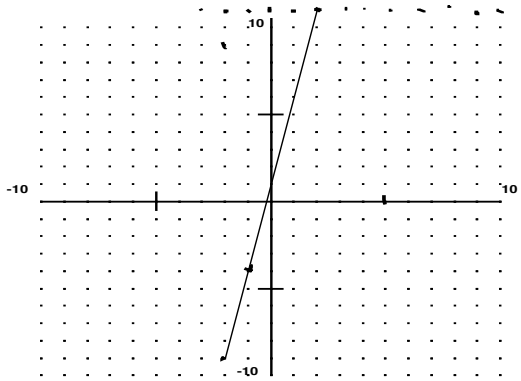

Figure 1. Green Globs graph of Y = 5X + 1

What happened? **The line is not really straight.**

211	**P**	Did it get thicker? Yes. Didn't it? No, wait! Then it didn't get... **It's getting bigger.**	Both are gesturing on the screen. P might be saying that it didn't get straight or that it got "bigger" and not "thicker" as she predicted.
212	**S**	I know. **It's getting bigger.** But how do we get that?	Agreement. Apparently now believes that prediction is wrong, correct answer is "bigger," but doesn't know how to justify this observation post hoc.
213	**P**	I don't know. Wait.Wait, wait, wait. (laugh) "Sketch..." What's sketch? "Sketch your prediction (S: Put it here...) on this empty graph and then try it on the	

[2]The Green Globs program and the experiment described here are the work of Susan Magidson, Judit Moskovich, and Alan Schonfeld. The term "green globs" refers to the dots connected when a line is drawn (Dugdale, S. and Kibbey, D. (1982). *Green globs and graphing equations.* Sunburst Communications. Pleasantville, New York). This educational game was designed to relate algebraic equations to their graphs. In the original configuration, students were given a graph with thirteen randomly placed points ("green globs") and asked to enter an equation that would pass through as many points as possible. A previous analysis (Clancey, 1993a) examined the first part of the Pam and Susanna interaction.

		computer." But what? "What happens?"	
214			S laughs
215	**S**	Reads "Sketch your prediction on this empty graph and then try it on the computer."	
216		(A brief exchange, inaudible. Laughter.)	
217	**P**	Not that, we can't do that. I don't know.	P rifles through sheets; S takes them.
218	**S**	Let's do it this way.	S possibly suggests filling in response after seeing the results.
219	**P**	We can't!	Possible reference to crossing out earlier answer. At lines 469 and 471 she says that they cheated.
309	**S**	It's just (inaudible) **It's not (P: it is...) a straight. (P: It is) See, it's not.**	S aligns worksheet plot shown above with ruler . Now is possibly saying "a straight"; before was clearly saying "as straight."
310	**P**	It *is!*	Oddly, perhaps because reference is to the drawn diagram and not the computer screen, P says line is straight for the first time.
311	**S**	Look. It's not a straight!	
312	**P**	*It is!*	(P takes the paper from S.)
313	**S**	See, it's over here. It's between this point and the other one.	S points to screen with pen. Apparently emphasizing that it is not on the Y axis?
314	**P**	It is. Wait. It *is*.	

Figure 2. S moves finger up the line Y =5X + 1, and says, "That's why this..."

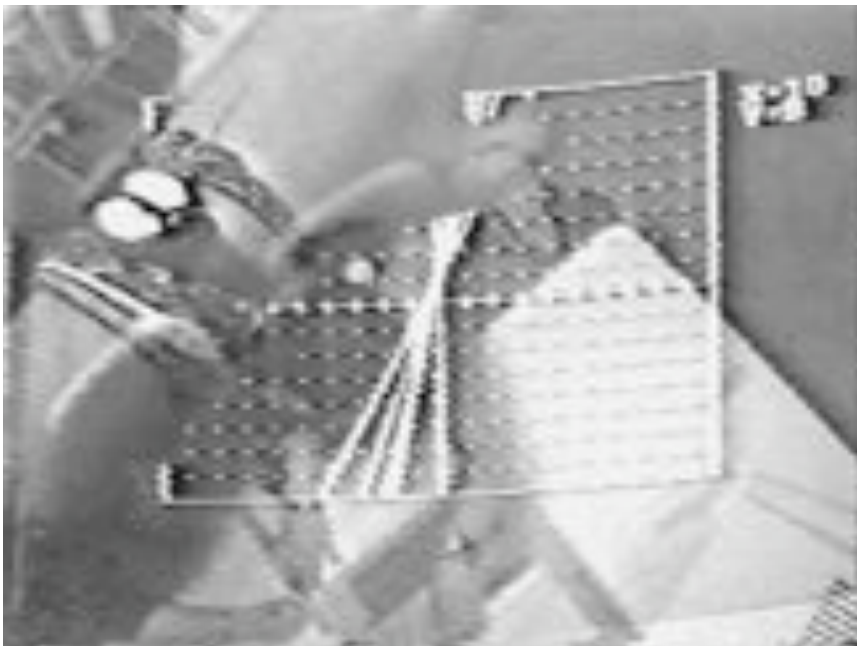

Figure 3. S continues gesture by indicating at top of screen how Y = 5X+1 is inside two others, with her thumb moving across, " ...not inside... that—"

315	**S**	It's not as straight. **That's why this...not inside... that—**	S gestures as shown in pictures.
316	**P**	Susanna, it's *straight!*	Very loud and certain.
317	**S**	Oh, but it...it'd have to be...it'd have...oh... well, it's not	Broken speech, difficulty articulating what she is seeing. Does not fit idea of vagueness as an uncertain choice between defined alternatives.
318	**P**	It is.	
319	**S**	It's not.	
320	**P**	It *is*. How can it—	
321	**S**	Okay, okay, all right.	Clearly resigns. P later refers back to this agreement?
322	**T**	Okay, it's straight.	Pleasant, somewhat humorous lilt. She's overhearing from the side.
323	**S**	(laughs) Okay, what happened?	Goes back and changes p. 5 prediction to add "not" and "straight" (difficult to reconcile with teacher's remark)
324	**P**	What happened? We got a (3, 9).	
325	**S**	It is...wrong. **The equation is going to get** thicker**.**	S says their prediction was wrong.
326	**P**	I don't know. **The equation is going to get** straighter**.**	Unsure tone. Insists on straighter.
327	**S**	No, the equation is going to get... **it's not going to get** straight**.**	Revised prediction. Insists not straight.
328	**P**	It's not?	
329	**S**	No.	
330	**P**	**The equation is**...(S: inaudible) **The equation is** not...	
331	**S**	**The equation is** going—	Chooses positive wording, disagreeing with P again.
332	**P**	**Is not going to get**—	P follows along with negative phrasing.
333	**S**	**Is not going to get** straight.	
334	**P**	That one. How could it...? (inaudible)	
335	**S**	It's not a straight.	Possibly says "**as** straight."
336	**P**	It's straight!	

337	**S**	Okay, okay. Put it's straight. (Inaudible.)	S gives in.
338	**P**	(Inaudible.)	
339	**S**	**Is this line straight?**	(spoken to teacher)
340	**P**	It *is!*	Note her willingness to answer so certainly, before hearing the teacher's response.
341	**S**	It's not.	
342	**T**	**What do you mean by straight? Do you mean, like, as in this is straight,** [gestures vertically on paper] **and that's not?** [gestures at 45 degree angle on paper]	Contrast T's presentation of alternatives with P and S's argument.
343	**S**	Yeah.	
344	**T**	So then, **it's getting close**, but it's not quite, is that what you're saying?	T gestures to the near the Y axis; uses the term "close"; she overheard S say this?
345	**S**	Mm hm. (turns to P) See? **So can we put here** [T & S laugh] **the line is not, is not really straight?**	
346	**T**	Mm hm. (long pause) I came in to tell you that you've been working for just about an hour. So if you've had enough and you'd like to stop, you can stop. If you'd like to work a little longer, you can work a little longer.	Acknowledges, meaning "you can write that." T exits after a discussion about continuing.
353	**P**	Yeah. **I think you said straight.**	P gestures with pen upwards, vertically on the page. Suggesting that S has contradicted her agreement (#337).
354	**S**	It's not straight, you know why? **Because it has to be like this with the points.**	S shows willingness to explain, more confident, apparently reflecting sense that T has confirmed her point of view.
355	**P**	I thought you said *que* straight; it wasn't a straight line.	P gestures to vertical.
356	**S**	**It is a straight line when you use a ruler,** but it's not when you do like this	S gestures with pen making points on paper; suggesting that freehand lines are crooked? Referring back to discussion with Teacher? Possibly "straight" means vertical (#55); "straight line" means not crooked (#38).
357	**P**	Real funny.	Tone is that S is indeed strange & humorous in her point of view.

What do you think will happen if you type in **Y = 1X +1**? **The line is going to get straighter.**
Sketch your prediction on this empty graph and then try it on the computer.

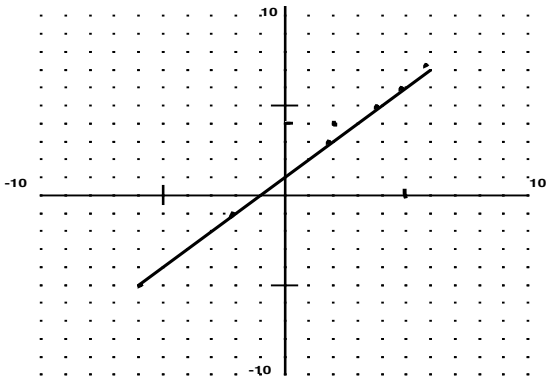

Figure 4. Green Globs graph of Y = 1X + 1

364	**P**	What do you think? You have to answer it first.	Again P insists on "not cheating."
365	**S**	Um...t**he line's gonna get straighter.**	
366	**P**	So write it.	Gestures to worksheet. Plays with pen in hair. Clearly not engaged.

367	**S**	Do another one with the same equation. Don't get mad!	Entering points on graph of Y = 1X + 1 on paper (above).
368	**P**	I'm not mad. Who said I was mad? There's not going to be (inaudible) right here.	
369	**S**	Try this one. **See, it's getting straighter.**	S notices that the line intersects grid cleanly?
370	**P**	No.	
371	**S**	You want to bet?	
372	**P**	**How weird, the lines get straighter...dots.**	P agrees now, apparently sees the line differently. Dots refers to grid.
408	**S**	It has to be 3...3...here. They have to be here. No, wait. Here. My fault. 5. **The line's getting straighter; I told you.**	
409	**P**	**Yeah... What do you mean, "straighter"? Like straight like the straight line right here?**	First serious attempt to negotiate a definition. Possibly echoing the teacher's question? P gestures along vertical axis, referring back to the teacher's interpretation.
410	**S**	Yeah. It is a straight, see?	Says "a straight," possibly meaning "a line."
411	**P**	It's straight, but if you put it like this...[turns the paper to 45 degrees]	Again, apparently repeats teacher's contrast; trying to show that being straight is not relative to the Y-axis, but a property of the line as an object?
412	**S**	Ha! (laughs)	
413	**P**	Isn't this vertical?	Suggests that the word for being aligned with the Y-axis is "vertical," perhaps to contrast with P's use of the term "straight."
414	**S**	Forget it. Forget it.	Forced to decide, S opts out.
415	**P**	What's horizontal? Like this [draws horizontal line on paper] or like that? [draws a vertical line on paper]	P pursues the point. Offers a clearer contrast about meaning of "vertical."
416	**S**	Horizontal's like this.. [draws a horizontal line]	
417	**P**	*Y* like this? [draws a vertical line]	
418	**S**	Alright. Let's just forget about it. **But it is straight.**	S holds her forehead, expressing frustration. Apparently refuses to acknowledge that Y-axis is called vertical and not straight. Possibly referring back to Y=1X+1 is straight.
419	**P**	If you say so.	Tone of resignation.
420	**P**	Wait...1...	Clears the screen
421	**P**	No! Yes. Huh.	
422	**S**	-1...	
423	**P**	Yes. Don't do any more!	
424	**S**	See! **See, it's getting straighter.**	
425	**P**	**Okay. Now, enough, enough.**	P gestures "stop" with her hand and puts her hand on her forehead
426	**S**	Wait, let me put it here.	Graphing Y = 1 X + 1 on computer
427	**P**	**All of them thing?**	Referring to the dots of the grid.
428	**S**	(Inaudible.) What...See? **It is a straight.**	Term "a straight" would suggest a property of the object, as in "a perpendicular"; S is noticing that there aren't any jagged segments?
429	**P**	**Oh, yeah, it is! I get it now. I understand what you're saying. I understand**.	Shifts point of view. Graph shows Y = X + 1, with the line clearly intersecting the grid (contrast with Y=5X+1 with grid points *near* the line).
430	**S**	Okay.	
431	**P**	**The dots, huh? The little dots, no?**	This is her second reference to the dots. Little dots refers to the grid?
432	**S**	Mm hm.	
433	**P**	Now I understand.	Apparently a true agreement.
434	**S**	Good.	

Were you right? Explain. **Yes, because the dots got on a straight line.**

{"Line" was replaced by "dots," which it written over several times.}

435	**P**	I'm sorry. "Were you right?" **Yes**. "Explain."	P touches S.
436	**S**	**Yes, because the line** got straight.	S focuses on a line, possible reference to pixels.
437	**P**	[writing] **Yes, because the line**...	
438	**S**	**got**	
439	**P**	**got** what?	Subsequent remarks suggest she wants to describe the dots, not the line.
440	**S**	**Straight!**	Humorous. After all this...
441	**P**	**The dots** ...were on... the straight line.	Speaks very slowly and deliberately. Says "the straight line" suggesting that the line was always straight? Salient feature is that the dots are *on* the line. Refers to grid?
442	**S**	Yeah. Erase that.	S agrees to mention "dots."
443	**P**	**The line...** **The line...** **The dots...**	
444	**S**	Yes, because **the dots...**	S uses the word "dots" for the first time.
445	**P**	**The line…** **The dots,** huh?	Acknowledges agreement that they are describing the dots.
446	**S**	No, **the line!** (laughs)	But S immediately rejects the focus on dots...
447	**P**	**The dots were**...	
448	**S**	Because I put here, "the line is going to get straight."	...because her prediction was about the line Y=1X+1, not the dots.
449	**P**	**The dots were on a straight...** **are on a straight line**	
450	**S**	Okay, okay. **The dots are in a straight line.**	S emphasizes "are" and "in"
451	P	**Were on a straight line.** **"Got?" Yes.**	**P emphasizes "were" and "on" while writing.**
452	**S**	Yeah.	
453	**P**	**The**... Hmmm! **The dots got on a straight line.** **The dots got on a straight line**. I understand it.	P reflects on the description and says that it makes sense.
454	**S**	(reading) "**Because the dots got on a straight line."** Isn't that...is that a sentence?	But S isn't sure she can parse it.
455	**P**	Yes. I guess.	
456	**S**	*Mira , tue eso* (mumbles).	*(Look, you are...*
457	**P**	*Callate*.	*(Shut up.)*
458	**S**	That's why Sofia...saw you (mumbles then laughs).	

Try some numbers greater than 5.
What do you think will happen as your numbers get larger?

462	**S**	Okay, 6...let's put 6.	S suggests Y=6X + 1.
463	**P**	No, we don't have to write anything. (laugh) (reading) "What do you think will happen as your numbers get larger?" **The...um...The...** [snaps fingers]**...(S: the line)...the dots are not going to be on that straight line.**	Possible reference to the 45 degree line they just drew. Almost a humorous delay in her pauses, as if trying to be funny... S interjects with line focus again; P contradicts immediately.
464	**S**	Okay, put that.	

Write down your prediction.
The line is not going to be straight.

Now try it on the computer. Were you right? Explain.
Yes, because the line was not straight.

470	**S**	Let's try it on the computer. "Were you right?" Yes.	
471	**P**	Yes. We cheated.	They simply wrote "not straight," nothing about the dots here.

Now let's try some small positive numbers (numbers between 0 and 1). What do you suppose these lines will look like?
The lines will look straight.

Let's try it. Clear the screen and type in these equations, one at a time:

Y = 1/2X + 1
Y = 1/3X + 1
Y = 1/4X + 1

What do you notice? **That the lines are not straight.**
What stays the same?

515	**P**	[reading with S] "What stays the same?" (they laugh) I don't understand.	
516	**S**	**What stays the same? The dots? No.**	S means that the grid is unchanged?
517	**P**	That they... Nope.	
518	**S**	"What stays the same?" The line....(P: no.) Let's skip that one"	

What changes? **The lines**

519	**P**	"What changes?" (both laugh) Ah, look, **these little lines** are getting...*asi mas* wide. (S: What?) Look, that thing are getting *mas*—	P gestures to the jagged segments with her pen. Trying to find another interpretation
520	**S**	**Wider? Thicker? Bigger?**	S brings back the earlier terms used to describe the jagged lines on the computer.
521	**P**	Uh huh.	
522	**S**	No, I don't think that's the correct answer.	S indicates that these descriptions aren't likely to be what the teacher intended.
523	**P**	(whispers) Okay now?	
524	**S**	I don't really know. I'm confused. Just— "what stays the same?"	Worksheet question is driving their interaction.

What do you think will happen if you type in **Y = 1/5X + 1**?
Sketch your prediction on this empty graph and then try it on the computer.
We think that the line is going to get closer to be straight.

Let's try some numbers smaller than 1/5. What do you think will happen as your numbers get smaller? Write down your prediction. **We think that the line is going to be straight.**

Now try it on the computer. Where you right? Explain.
Yes, the lines didn't get any closer to be straight.

532	**S**	The lines—	S types Y = 1/5X + 1 again
533	**P**	**I told you that they're getting bigger.** **The lines will get closer to the straight line.**	P refers back to claim that they the "little lines" are getting bigger. But she describes lines as getting closer to the X-axis. The progression of equations suggests focusing on what's different rather than the common Y intercept.
534	**S**	To the what?	S says "the."
535	**P**	**To a straight line.**	P says "a."
536	**S**	Yeah.	They agree from here to the end because P drops her observation about the little lines getting wider.
537	**P**	Write it down.	
538	**S**	It doesn't (inaudible)	
539	**P**	"What do you think will happen"—	
540	**S**	Okay. I think—or we think. We think. We think that the line is going to, is going to get, we think the line is going to get close to being—	S changes "I" to "we" in worksheet. Manifests her sense that they agree. Says "close" again.
541	**P**	to be a...	emphasizes "be"
542	**S**	**—to be a straight line**	possibly no "a"
543	**P**	We think that that is going...?	
544	**S**	We think that we, that the, that that is... we think that that that is going... (laughter) **I have no idea what I'm thinking.** (Laughs.)	S struggles with wording. Indicates that she doesn't understand the lesson.
545	**P**	(Laughs.) Oh, God.	P looks up to the ceiling
546	**S**	Do you want to go now? It's four fifteen.	
563	**P**	Call her.	
564	**S**	Let's get out of here. Do you know where the garbage can is?	

ANALYSIS OF THE GREEN GLOBS INTERACTION

What is happening here? First, the students were never told what features to look for in the graphed lines, simply to compare them. The text opens by using the word "straight" twice: "these equations are straight lines... will produce a straight line..." So what is a non-straight line? Told that straightness is a property of some equations, perhaps the students can discover equations producing lines that aren't straight?

P and S's interaction can be studied from many perspectives. Some of the questions to consider:

- What do they experience that's difficult to describe?
- What is their *practice of description*? Why and how are new terms introduced, shared, and written?
- What are the differences between the students in adopting different views, seeking and giving explanations, promoting collaborative interaction, introducing new terms, adhering to the worksheet, gesturing, etc.
- How do the students assess their own understanding?
- How do different modalities (nomenclature, diagrams, instructions, computer graphics, gestures) relate in the activity to foster understanding?

- Are the dynamics of their interaction confirming each other's perceptions or leading to opposition? Why, if one student "gets it," do they end up with different interpretations?
- What are the interactions between the different aspects of activity? Social-interactional ("what I'm part of"); representational-cognitive ("what I'm inferring and planning"); operational-behavioral ("what I'm doing here and now") (after Leont'ev, 1979).
- How are they generating equations to graph? Are they testing hypotheses?
- How does the worksheet's design direct, help, or inhibit their understanding?

A conventional analysis might focus on the logical argumentation, evident for example in lines 411-418 where P confronts S with two defined alternatives. But my interest here is in *conceptual transformations*, aspects of non-descriptive understanding in images, gestures, and emotion, which are dialectically developing with the spoken and written descriptions. (By dialectic, I mean that conceiving—a neurological process—and describing causally influence each other.) I am especially interested in development of sequences (e.g., as full sentences are written in the worksheet (325-333)), repetition in behavior sequences ("it is"—"it is not straight"), and means by which the student resolve impasses (e.g., dismissing, laughing, shifting levels).

I will begin with a summary of the interaction, and then consider in turn aspects of perception, reference, description, collaboration, and breakdown. I conclude the analysis by considering how describing relates to conceptual coordination.

Summary of interpretations

P and S give many explanations for why the lines are straight or not. Besides S's belief that a vertical line is straight, both P and S notice that lines between 45 and 90 degrees are jagged, an effect caused by coarseness of the screen display (number of pixels per inch). Early on S also notices that their hand-drawn lines are not straight. At the very end, P notices that *the equation* Y = 1X + 1 intersects the dots of the grid (which evidently S agrees is straight because the jagged segments disappear).

The following aspects of the interaction are of special interest:

- P's understanding shifts between jagged segments and alignment to the grid. S appears to shift between three interpretations, holding most firmly to verticality.
- S appears to use the word "straight" in two ways at the same time, corresponding to "being in a line" (definition of straight line) and "being vertical" (a kind of straight line) (418).
- S's insistence on "the dots" near the end apparently prompts P to look again; she sees "the little dots" as being relevant for first time (427-431).
- P introduces the terms "thicker" (133[3]), "width" (171), "bigger" (195), "dots" (372), "vertical" (413), "horizontal" (415), "little lines" (519). S does not introduce new terms (indeed, she cannot remember the word "vertical").
- S's descriptions are qualifications on "straight"—"always straight" (116), "stay straight" (142), "straighter" (147), "not very straight" (194), "not as straight" (195), "not a straight" (309), "going to get straight" (326), "getting straighter" (369), "got straight" (441). S's qualifications are relative to an ideal reference (being in a line or the vertical) and consistently mention change in appearance.

[3]Some quotations appear early than the transcript reproduced in this paper. Line numbers are shown for reference, to indicate relatively where they appeared in the interaction.

- P's explains by naming a particular configuration ("straight like the straight line here?" (409)) and mimics the teacher (342) by drawing vertical and horizontal lines to illustrate her meanings (417-418). She attempts to disambiguate S's meaning.
- S explains by classifying an instance as a member of a category ("it's not a line" (38)), pointing to a visual property ("it has to be like this with the points" (353)), and describing an action ("when you do like this" (353)). She shows P what she means, but doesn't confront her with choices.
- S is tuned to the requirements of the teacher and the worksheet (212), what is a plausible response (325, 522), and the timing of the session (546). P appears more oriented to understanding S.

Perceptual reorganization

What symbols are *given* in this problem? In some sense, only pixels are given. If the input were obviously straight lines, as intended by the teacher, P and S would only be "symbolic information processors." Rather, what they see, understand, and how they talk arise together, codetermining each other. My objective in what follows is to provide evidence for this claim, and show how visual categorizing, referring, and describing develop together.

In Bamberger's (1991) terms, the question for the children is what *figure* should be seen as straight? The process is not simply classifying lines, but constructing configurations into objects, by which the word "straight" can be given an interpretation. These configurations are perceived at different times during the process of making sense of each other and the worksheet. They include: a thick bundle of lines, jagged "little lines" (519), and little dots (the grid).

In itself, referring to "the little dots" is a figure-ground shift. The grid is now no longer just a background, but perceived as objects to be described (the little dots get lined up)—there is shift between "in-ness" and "on-ness" when viewing the display as a configuration. In effect, the students agree that straightness has something to do with "being lined up," but it's unclear what gets lined up with what. Figure 5 summarizes some interpretations.

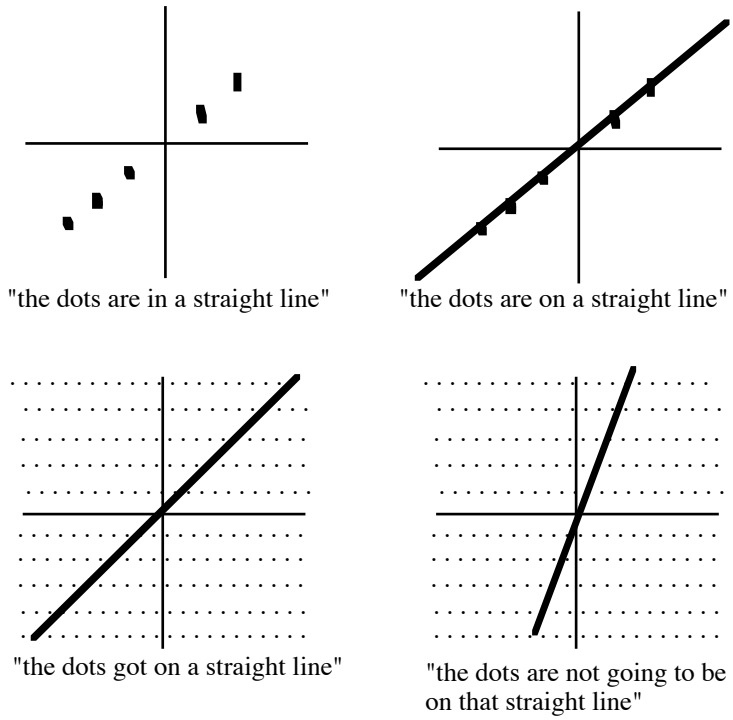

Figure 5. Graphic interpretations of what the children are seeing (the forms they are perceiving and relating) when they describe alignment (see transcript lines 441/453 and 448/450)

New interpretations are based on perceptual regrouping: In considering how the points of the equation are lined up, the dots of the grid are irrelevant. But when we include the dots of the grid, a previously "straight line" is no longer "lined up." Thus, different ways of talking are grounded in *different images.* The meaning of "straight" depends on what objects are perceived and grouped in the scene and the relations of these objects to each other. Significantly, an "object" may be a space between two lines (Figure 3).

In traditional schema theory, the meaning of new terms given by a teacher is defined in terms of old terms by generalization and specialization, as well as by correspondence to a given, stable scene in the world. But because the vast majority of symbolic models (such as ZBIE) do not engage in visual processing, researchers didn't acknowledge the perceptual reorganization that might be involved in learning new meanings.

The construction of reference

To carry this analysis further, let's consider what the students are conceiving when they are gesturing and using indexicals like "it," "this," and "that." A classic example is shown in Figures 3 and 4, when S says, "It's not as straight. That's why this...not inside... that—" (315).

In general, perception is occurring on several levels: where to look and what constitutes an object, which is the topic of description (the reference of "this" or "it's"). Once directed to an area by a gesture, one still must know what level of detail is relevant. For example, P says "the little dots" (431) and "these little lines" (519) to refer to the grid and the jagged segments, in contrast with the plotted dots and the plotted equations. But now, in seeing some figure, one is *conceiving of a difference.* This difference is not a thing in itself, but a change or a contrast (Bateson, 1972; Roberts, 1986, 1993).

The students describe their conceptions by calling attention to a figure, some particular object or configuration of objects, which they describe by contrasting it with the surroundings in terms of temporal change, location, shape, and internal configuration. Indeed, these contrasts are multidimensional, as shown by Table 2.

Table 2. Varieties of difference or contrast described in explaining conceptions of "straight line."

	Meaning of "straight"	**temporal change**	**internal relation**
location	vertical or aligned to the grid	"straighter" "closer"	"on the straight line" "not inside that"
form	not bent	"thicker" "wider"	"little dots" "little lines"

Three meanings of "straight line" are described by contrasting a figure in terms of location, form, change over time, and internal relations (e.g., the little lines that make up the big line). Such descriptions, combined with gestures in a shared visual space, enable the children to co-construct figures. These figures (ranging from a bundle of lines which is getting thicker to the dots of the grid) constitute a particular detail which for the children is *new*, a basis for reconceiving the meaning of straight. Put another way, conceptualization is visible in this interaction as the description of changes and contrasts the children perceive. Their describing and gesturing first acts to separate figure from ground, and secondarily to define "straight" in terms of the contrast they perceive.

Indeed, the difficulty of the interaction between P and S appears to lie in S's proclivity to describe contrasts, and P's proclivity to describe figures. P must then work to understand what S perceives is changing (since S says "this" "that" is changing, getting closer, getting on, etc.). In general, S adopts P's contrast words ("thicker" "wider") but never introduces her own names for focus-details.[4] The following exchange is typical:

170	**S**	The line gets...
171	**P**	**The length**. No. **The width**.
172	**S**	Gets...
173	**P**	...width
174	**S**	**Width gets thicker**
175	**P**	Yeah
176	**P**	The lines get thicker?

Indeed, every single reference to "it" or "this" in S's discourse refers to "the line," "the lines," or "the equation" (which refers to the graphical form on the screen). S resists or rejects

[4] By focus-detail, I mean perceptual figures that are being incorporated in a description such as dots and lines and their perceived attributes; the term is used by Bartlett, 1932 in characterizing story-telling when remembering.

every single attempt to enter a statement in the worksheet in which the subject is something else, including the dots, little lines, or bundle. When P nevertheless enters such statements, S accepts them passively.

Given that the worksheet is calling for the children to say that the slope or angle of the lines and the Y-intercept is changing, this resistance to adopt another focus-detail is a fundamental problem for S. P is wandering around looking for another figure by shifting grainsize and making a figure-ground shift; S is fixated on describing the lines as wholes. We might say that S's preferred contrast, that the lines are getting closer to the vertical, is "conceptually close" to the idea that the angles are changing. But S's contrast is with respect to a fixed reference, and she is focusing on the *ends of the lines*. P and S are simply not looking at what is occurring in the *spaces between the lines*, down near the origin. They need to *see the space as a figure (*which we call an "angle"); this perceptual reorganization never occurs.

To reiterate, understanding that all the lines on the screen are straight does not mean merely relating some concept description to "instances" which are on the screen. A descriptive model of perception views seeing and recognizing as this kind of feature matching and mapping (e.g., see Larkin & Simon, 1987). Roberts refers to this as "reference *qua member*," in which some figure is claimed to be an instance of a general thing ("this is a straight line") or to have a property of a type ("this line, like all lines, is straight"). We can find examples of such descriptions in P and S's interaction, such as when S says, "It's always straight" (117).

But Roberts points out that the construction of reference occurs as "reference *qua particular*," in which a figure is described via a contrast as a thing. By this view, seeing and recognizing is a process of *creating features* and inherently involves visual reorganization (Clancey, 1997). The concept of the general thing is then developed dialectically, by the inclusion of this example. Describing within the process of learning is not just pointing, naming, and defining, but separating something out from the background and describing the figure *as a contrast*. That is, the figure is not something that stands alone, but is only *known as a difference* over time, of form, or within a larger configuration.

As Bateson emphasized, the contrast and hence the visual concepts are not located in particular things. The reference is not to an object per se, but to a difference: The lines are getting closer, wider, thicker. "The dots got on a straight line"; "the little lines are getting wider." A contrast is an experience, occurring over time or within a process of looking. A visual contrast is within a particular area, but is not a property of a particular thing in isolation. In saying that the lines are getting "closer," for example, S focuses on the difference in distance of the line (segments) from the vertical; in gesturing (Figure 3) S shows us that she perceives this space as a figure, for which the lines are now a ground.

A ground is also a visual conception. In 316 S is attempting to describe how the figure she is currently seeing (referred to by "it") is different from another visual conception—how she sees the meaning of "straight." When the ability to put her experience in words fails—she plainly knows that a contrast exists—she falls back on denying the applicability of P's description. The conversation then becomes a shouting match.

Finally, referring back to Table 2, the dimensional analysis of the multiple interpretations of "straight line" suggests that S is able to hold to both "vertical" and "not bent" because they arise from different visual organizers—one conceiving difference in location and the other difference in form. S appears to smoothly move from seeing the hand-drawn lines as not straight (101) to seeing the thickness (jaggedness?) of the lines as not straight. But she is shifting her

point of view when she agrees with the teacher that the Y-axis is straight (342)—a matter of location and not form.

At the end, when they graph Y = 1X + 1, S is perhaps still seeing the line in terms of form—she had previously indicated that smaller numbers (with slope approaching 45 degrees) the line is getting straighter (147). Ironically, the salience of intersection with the grid leads P to see this line as straight, but for a different reason, one based on *location*, not form. Hence P and S can agree on the description, a 45 degree line is straight, but are unable to agree on the definition because they are contrasting their experience on different dimensions.

Therefore, the inability to communicate stems from not articulating these different ways of organizing the visual field. S never acknowledges that she is viewing "straight" as a matter of form at one time and location at another time. This same difference is replicated in P and S's conception at a given time, and exacerbated by S's tendency to merely agree when P says something she apparently doesn't understand (432).

The analysis of the relation of description and visual conception of contrast is fundamental for understanding the interaction of P and S. However, there are many subtle aspects of how these descriptions are created, through the interaction of different modalities. In particular, the children are coordinating different actions (speaking, drawing) and coordinating their own interaction (filling in the worksheet, taking turns). In subsequent sections, we will consider how the conceptualization we have just described occurs within larger, coordinating frameworks with serve to constrain and supply resources for what the children see and say.

Aspects of representing

Broadly speaking, representing includes much more than describing or drawing. Besides names and phrases, P and S are creating other forms which are intended to represent their understanding of the worksheet's questions and what is happening on the display screen. Representing, as an activity, involves a number of different activities using different modalities:

- spoken phrases ("the little dots")
- drawing ("*Y* like this?" (417))
- gesturing (Figure 3)
- calling attention to details ("Look! Which one is that?" (167)) and looking again to verify a hypothesis (211)
- following worksheet directives (203)
- focusing a written response on some figure (443-447)
- commenting on a partner's representation ("No! The lines are getting bigger." (195))
- abstracting perceptual patterns ("Ah, look, these little lines are getting...asi mas wide..." (519))

These activities are not all descriptive in origin and nature. As Bamberger and Schön (1991) emphasize, the students are engaged in a "conversation with materials." They are taking turns writing and interacting with the computer; they are choosing equations to display; they are coordinating ways of understanding each other as they are following the worksheet's directives. As Dewey (1902) would emphasize, P is reinterpreting and *looking* for other features. She is attempting to reconcile her view that the lines are not straight (116) with the worksheet, the teacher (344), and S's disagreements. Probably she would not have come up with ideas about the little dots and little lines if S had agreed with her at the end.

The symbolic approach claims that all these actions are driven by subconscious descriptions. But S's stumbling and repeated references to "this" and "that" as she points suggest that she has no words for what she is seeing, the figures and contrasts that are of interest to her. At a basic level, we see that *the work of representing* is adapting and learning words that describe personal experience.

Collaboration

At the beginning, the teacher said "I'd like you both to work together, so I'd like you to take turns typing and take turns writing." Indeed, their sharing is remarkable. On the other hand, the obligation to fill in the worksheet prompts heated discussion and sometimes forced resignation. Of special interest is how they suggest words for the worksheet, which for the observer conveniently reveals their different focus of attention ("the dots" "the line") (Garrod & Anderson, 1987).

Although the worksheet is the product of two people collaborating, the students have a marked sense of independence. Beside the obvious banter— "it is" "it is not"—P&S have clearly indicated their identities on the worksheet by how they dot the letter "i." Furthermore, their activities are oriented in different ways. S assumes responsibility for satisfying the worksheet: She is the only one who asks questions of the teacher; she points out inconsistencies between what they observe and what they predicted; she turns to the worksheet more often for direction ("Okay, what happened?" (323); "what stays the same?" (524)). P worries a bit more about reaching a shared understanding: she deliberately seeks explanations for S's claims, both by reinterpreting the display and presenting choices to S to interpret.

The high degree of interactively is most pronounced in their typing and exchanges when filling in words (142-153; 169-186 and 437-449). An example of how P and S type an equation (Y = -5/3 X + 7) appears in the following transcript.

Line		**Statement**	**Gestures**
79	**P**		P types Y =
80	**S**		S types 5 / 3
81	**P**	Wait! It's negative.	brushes S away from keyboard; erases 5 / 3; types - 5
82	**S**		S types /; P reaches forward; S types 3; P reaches forward; S erases 3
83	**S**	X there.	gestures to X
84	**P**		P types 3 X; S gestures to +; P has both hands on keyboard; P types =; erases =; types + 6 . .; erases .; types 7; S gestures to return; P presses return.
85	**S**	Yeah.	

But there are as many examples of discord. On several occasions, the children give up in their attempt to work together, but allow the partner to proceed with her preferred action. S gives in by saying "Okay, okay" (321, 337, 450). P gives in by telling S to write the statement she prefers: "Write it!" (197, 366). Both children express exasperation at different times, looking up at the ceiling and holding their heads in their hands (418-425).

Learning may be collective and based on communication and interaction. But nobody can learn for somebody else.[5] At a certain point, each child expresses her individuality and present understanding by calling a halt to an exchange and moving on. Thus the task of filling in the worksheet is satisfied, and the individual's sense of personal understanding is preserved—at the expense of not understanding the partner and not representing for the experimenter-teacher what each person understands.

Breakdown, action and talk

Disagreements occur when different conceptualizations lead to different ways of describing. Impasses result not only because P and S can't agree what to call something (or the correct syntax), but because as I have discussed above, they see different figures or conceive of the figures in different ways.

At different times, each child finds that she cannot continue the activity because she is unable to resolve this conflict. When these discussions are not focused around how to fill in the worksheet, but are instead squarely about their concepts, a breakdown may be resolved by simply calling a halt to the activity. Perhaps the best example is when P is leading S through the definition of vertical and horizontal, which S halts by saying "Alright. Let's just forget about it." (418) At other times, a breakdown may be resolved by asking the partner to give her more time (P says, "Wait." (142-143)) or by marking an inability to understand (P says, "How weird." (372)).

An impasse is a *discoordination*, a breakdown between how ways of seeing, conceiving, and talking are dynamically related. Experience of an impasse is often accompanied by an emotion or attitude, that Bartlett (1932) emphasizes accompanies a new orientation. Because a person must be experiencing *something*, the emotion appears to substitute for the previous (ineffective) conceptualization. For example, in 309-320 the girls contradict each other six times, interposed with requests to look and explanations. S finally breaks out by simply saying "Okay, Okay, all right." The teacher breaks in at this moment with a humorous lilt, "Okay, it's straight" and then S laughs.

Here is a summary of how impasses are resolved in P and S's interaction:

- Try to control or end the offending event. In the face of S's insistence that the 45 degree line is getting straighter, P says, "Don't do any more!" (423) to get S to stop plotting points.
- Laugh or dismiss the behavior, "Real funny" (357).
- Ask for a justification, "I know… but how do we get that?" (212).
- Appeal to logic and authority, by referring to the worksheet's directives, their previous responses, or the teacher, "I don't think that's the correct answer" (522).
- Request a clarification of the reference: P: "It's straight." S: "Where?" (126).

[5]Thanks to Janni Nielsen.

- Suggest a rephrasing for clarification of the figure and contrast: S: "Width gets thicker." P: "The lines get thicker?" (176).
- Force the partner to look again, accompanied by gestures and descriptions of focus-details, "Look. It's not straight!" (311). (See also the gestured explanations in 211, 313, 315, 342, 356, 409, 411-417, 519.)
- Wait for later resolution and move on with a simple acknowledgment: S: "The little dots, no?" P: "Mm hm." (453).
- Classify the activity as irrelevant: S says, "Forget it...let's just forget about it" (414, 418).
- Move to a larger coordinating conceptualization of the relationship: S tells P not to get mad (367) and reminds P of an incident with a friend (458).

Impasses may lead to new conceptualizations for coordinating the activity. For example, after a point it becomes apparent to P and S that their predictions of how lines will appear on the screen do not fit their perceptions of what is happening. Consequently, they must adjust their conceptualization of their activity: Accept that some of their work is wrong or "cheat" by looking at the result on the computer first (or erase previous responses). Thus, the procedure for how to fill in the worksheet develops in the course of the activity. This is what I mean by a new coordination. The disagreements are not rooted in just the definition of "straight" but in differing conceptions about how the experimental session is to be carried out.

Coordinating multiple interpretations in two languages

To understand how concepts and words are related, we must consider the possibility that S is using a single English name ("straight") for several concepts she associates with different Spanish words, her first language. The effort to coordinate a single word with multiple concepts, and hence multiple ways of seeing, may be the cause of her difficulties.[6]

In particular, S may be conceiving of "straight" in the everyday sense of "derecha" (una linea derecha), which is the opposite of crooked or oblique. This meaning incorporates the sense of standing up straight (i.e., vertically), straightening a picture that is askew (i.e., making it perpendicular to the floor), driving straight through the city (i.e., in the same direction), and so on. In this sense, the Y-axis, the non-jagged computer lines, and the lines drawn on paper with a ruler are straight because they are *derecha*.

Spanish, unlike English, uses a different word for the mathematical sense of straight, "recta" (una linea recta). Repeatedly, S refers to "a straight" (309, 311, 335, 410, 428), suggesting that she knows this meaning, too.[7] In this sense, all lines are straight (117), so the word "line" in "a straight line" is redundant.

Furthermore, "recto" in Spanish means both "straight" and "right" (as in rectángulo). In English, we use the Latin "rect-" prefix without realizing the double meaning ("rectilinear" means "straight-lined" but "rectangular" means "right-angled"). In this sense, one may say that the Y axis is a paradigmatic *recto*, for it indicates the right angle in the graph (90 degrees) and is a straight line.

---

[6] I am indebted to Sue Magidson and Judit Moskovitch for the initial analysis which appears in this section.

[7] Listening to the video tape, we are faced with a perceptual problem. One can almost shift between hearing S say "not as straight" and "not a straight" at will. However, I found five occasions where the phrase sounded more like "a straight," and no occasions where "as straight" appeared more likely.

So when S says at the very beginning, referring to the drawing on paper, "It's not a line" (38), she probably means it's not *una linea recta* because it's not *derecha.* When she says, "It's always straight" (117), referring to the lines graphed by the computer, she means that they (*linea recta*) are always *derecha.* In both cases, *derecha* is conceived as an inherent property of *una linea recta*. But when P shows that the lines are getting thicker, S shifts to viewing *derecha* as a *description of appearance*, which may or may not apply to a given *linea recta*. Indeed, when first describing this appearance she shifts from saying "the line" to "the *lista*," meaning "strip" (179), a figure on the screen. Significantly, she immediately afterwards refers to this figure as "the equation" (207, 325-331), showing lack of distinction between the mathematical entity and what she sees on the screen. At this point, her understanding appears to be that some *equations* actually do have the property of being thicker than others.

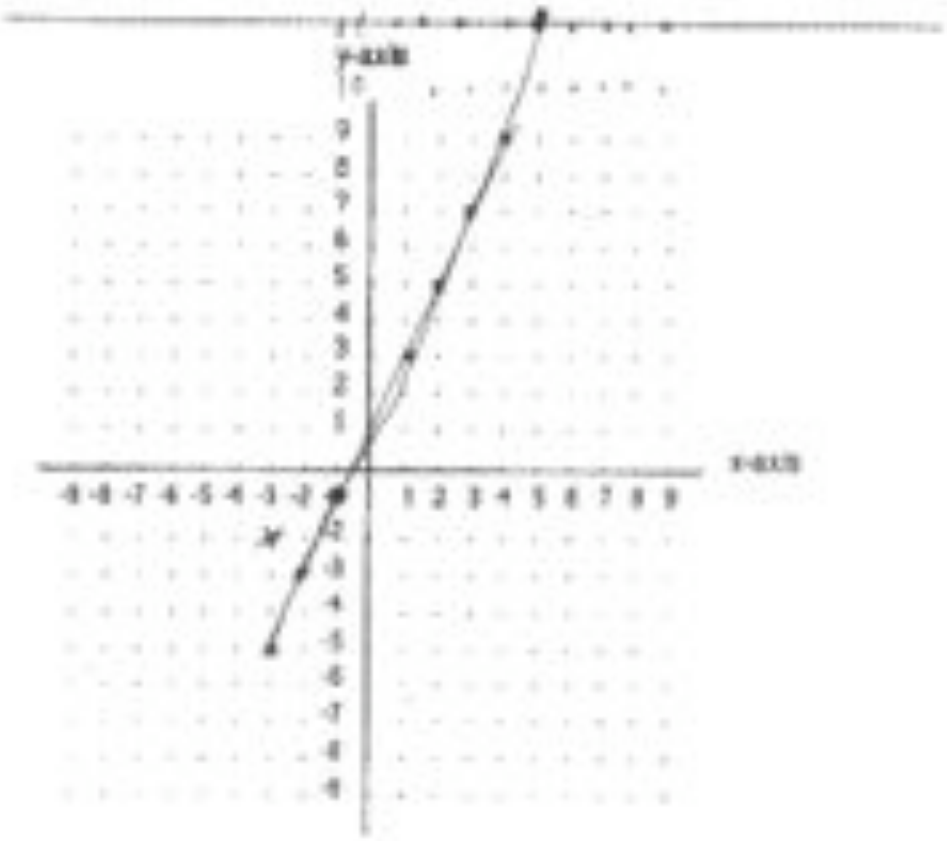

Figure 6. Actual plot produced by P and S with teacher's help.
Notice that freehand drawing of the line is not straight.

It is difficult to tie S's interpretations into a neat bundle. The evidence suggests that S shifts between interpretations of "straight" as she conceives of different meanings corresponding to different Spanish words. She appears to know the meaning of *una linea recta* in Spanish, given her repeated use of the phrase "a straight." Her sense of redundancy in the English phrase "a straight line" may have brought the contrast between *derecha* and *recta* into the foreground: some lines do not appear straight. Given the messiness of their initial drawing (Figure 6) and P's subsequent claims about thickness, S conceives the lesson as characterizing which *recta* (equations) are not *derecha* and why.

Coordinating acts of redescription

I have listed multimodal aspects in the activity of representing, but haven't considered how these are brought together. First, following Dewey (1896), we would want to explore the hypothesis that conceptualizing, experiencing, talking, and manipulating are occurring as one coordination. That is, these aren't independently occurring processes or variables, but arise as an interactive product. Describing is not just saying or expressing something in words, but one observable aspect of the process of coordinating activity. We observe these recoordinations when the children *recast previous descriptions*, both in looking back at a response and in adopting and modifying each other's terms and phrases.

The traditional approach of modeling concepts as networks of words would suggest that S has a subconscious description of what she is seeing. According to the symbolic view, S is not very good at explaining her reasoning—the meanings are organized inside, but perhaps are inaccessible or inconsistent. A more parsimonious explanation is that nonverbal experience is organizing S's action, and she simply does not have any words to describe what she is seeing.

S's stumbling is not evidence of "reading out" or deductive inference. This is *what representing for the first time looks like* (540): "We think that we, that the, that that is... we think that that that is going... (laughter). I have no idea what I'm thinking (laughs)."

The work of recoordinating visual conception in comprehending and recasting descriptions is evident in the alternatives offered by the children: in/on, the/a, the dots/the line, being/be, thicker/width. Choosing a term is choosing a point of view—n*ot describing what is already seen*, and certainly not merely translating what is already described.

In contrast with the kind of chronological sequencing in music learning discovered by Bamberger, language affords more reordering. Words can be rearranged more easily than the body. Reorderings themselves change the meaning of the parts (a dialectic effect). Just as a sequence in a melody defines the hearing of the individual tones, rearranging words is creating a configuration, within which the parts relate meaningfully. Just as individual tones can't be "recovered" unchanged, the children aren't manipulating words (atomic meanings) that have some fixed (atomic) meaning independent of their use. A rephrasing is an act of constructing a meaning.

The students' striking turn-taking in typing frequently occurs when they are composing sentences. Several aspects of conceptual coordination of sequences are manifest:

- Use of anchors, such that phrases are repeated from an accepted head (i.e., repeating what has been agreed)
- Incorporation of perceptual details, i.e., agreeing what needs to be described and what are the defining perceptual characteristics of an abstract mathematical definition

This phenomena is strikingly evident in the transcript. In the sequence 325-329 only an adjective is at issue. Notice how they keep repeating the sentence from the beginning and then "not" becomes a focus-detail. In the next sequence (330-333) the anchor shifts from the subject to the verb. Notice how both students incrementally add to the sequence and how P twice introduces "not," but S persists, producing a result identical to 327. Sequence 435-454 shows the problem of agreeing on a focus-detail, a figure, which needs to be described. In terms of conceptually coordinating sentence construction, the problem is to agree on a subject anchor—is the topic the dots or the line?

The incremental constructions reveal negotiating about the subject (dots vs. the line). Interaction 448 brings the two girls back exactly to 441, with P's statement, but now "the dots"

has been fixed and the focus shifts to the verb, "were" vs. "got." S had introduced "got" (438); P acknowledges this (451) and then accepts it. But the result is ultimately P's view (contrast 454 with 436); for S it isn't clear that this is even a sentence, let alone correct.

Again, my interest here is to go beyond simply observing that turn-taking occurs, to examine the *sequence* as revealing *conceptualizations* that organize the interaction and conceptual transformations occurring within it. In particular, the above sequence shows repeated questioning about the subject and verb, as both girls are directly involved in constructing a written sentence. Although P's viewpoint dominates (she is writing), she incorporates S's verb. So the apparent individual choices (P: the dots were; S: the line got) become composed as "the dots got."

An attempt to construct a meaningful statement may also fail (see 515-518). Here, as in 441-444, P and S both make bids for a figure—"that they" and "the line." P interrupts S ("no") and S acknowledges the evaluation. So they decide to skip this question. In effect, the children are unable to coordinate *some* meaning of "same" because neither can conceive of an appropriate figure.

To summarize, to understand conceptual change we cannot assume that problems are merely texts and diagrams. The children's problems consist of much more than comprehending text; impasses are not merely matters of understanding a referent of an already conventional representation. To understand perception, we must not assume that the world is given as objects with inherent properties. To understand the nature of description, we must not assume that concepts are named and described properties stored in memory.

CONCLUSIONS: LESSONS LEARNED FOR "INTELLIGENT TUTORING SYSTEMS"

The sometimes confusing interaction between P and S is sobering for designers of computer interfaces and instructional text. Clearly, more guidance about what to look for on the screen would have been possible and might have helped. But in practice it is impossible to anticipate all the alternative ways of seeing the screen. Understanding what "straight" means is not a matter of memorizing a definition, but of coordinating (and creating) possible meanings of the words with what you are seeing. For example, suppose we told Paula and Susanna that "straight means that the dots you plotted are lined up." What does "lined up" mean? Do "the dots you plotted" include the intermediate dots the computer filled in for you, that is, the pixels you caused to appear on the screen? I would hope that P and S's interaction would dismiss any designer's assumptions about simply engineering the system to avoid student misconceptions.

To understand better the student's point of view, we must focus on how people create representations, perceive symbols, and attribute meaning in physical manipulation of materials. We begin with a new contrast: *Teaching a pre-formalized curriculum vs. studying how a new language develops.* Attempting to relate levels of analysis—perceptual, deliberative, and social—leads us to reconceive the nature of misconceptions, as well as the resources enabling successful learning. Successful design does not depend on only—and ultimately cannot rely upon—careful choice of words and diagrams. The realization that *contrasts*, perceived in experiences over time, is the source of new conceptualizations provides a fundamental shift in how we view lesson planning. Although, it has been known for sometime that ordering lessons is important, the focus has generally been on logical prerequisites, based on the idea of composition and refinement of descriptions. To step out of this "representational flatland," we must understand learning as a process of multimodal recoordination during interaction with physical

materials. That is, we must develop our lessons around the shift in figure-contrast which occurs as meaning is constructed. Such a reframing of the learning problem may indeed help us to consolidate arguments about the many methods of instruction—coaching, discovery, tutoring—which are otherwise viewed as competing alternatives.

Equating human knowledge with descriptions (e.g., expert system rules), eliminated the grounds and origin of belief, and greatly oversimplified the complex processes of coordinating perception and action. Put simply, a learner participates in the creation of *what is to be represented* and *what constitutes a representation*. This dialectic process can be modeled by schema transformations of assimilation, refinement, etc. (Norman, 1982), in which descriptions are logically combined in an individual mind. But such a mechanism posits a set of descriptive primitives out of which all expressions are formed. The analysis here suggests that although these primitives may exist, they are so general and open to reconfiguration (as in figure-ground shifts) that an additional theory is required to explain how such primitives are configured to form a visual conceptualization.

Specifically, a theory based on mere recombination of primitives mapping to a "perceptually obvious" world (Larking & Simon, 1987) doesn't explain how new representational languages are created or conventional notations are learned. The analysis shown here suggests that it is insufficient to posit that descriptions are controlling how visual primitives are assembled; rather the learner's common experience is that she sees figures on the basis of contrasts she cannot yet describe. A mechanism grounded in descriptions and visual primitives also fails to account for individual differences, because it assumes that there is one objective world of features that everyone can perceive. In short, the exclusively symbolic approach fails to acknowledge or explain what is problematic to the learner, namely determining what needs to be understood.

But what kind of cognitive theory do we need to exploit communication technology? Have we been correct to assume that pedagogy must be grounded in an accurate psychological model of knowledge, memory, and learning?

Ironically, the same constraints that made ZBIE appear successful may be employed to some degree in a computer instructional system: One may offer a predefined list of descriptions in menus and operations, which channel the student into the terminology and distinctions of value in the coordinate system being taught. For example, rather than a freeform workbook, in which students write responses, what if P and S had been given multiple choices such as "the line is closer to the Y-axis (vertical)" and "the line is closer to the X-axis (horizontal)"? Suppose that this were hypertext, so the students could select terms like X-axis for further information. By having examined a variety of student responses, such as those by P and S, the designers could anticipate a broad range of difficulties (not only misconceptions), and thus craft a flexible system.

I believe better engineering has merit and could generate a more productive interaction than the handwritten worksheet. Of course some caveats are mandatory: First, there is no guarantee that such a system would work for all students. Second, the problem of modeling the student's understanding and offering assistance on that basis is finessed. And this second point is what bears some discussion.

Here are the points I take to be most salient:

- A human teacher cannot expect to follow, understand, and correct all aspects of a student's behavior.

- Even after dozens of hours of analysis over the course of a decade, I have not fully understood what P and S are experiencing and doing.
- P and S find their own individual experience problematic, uncertain, and frequently at loose ends. That is, they are lost.
- At key junctures, P and S's behavior appears to be ill-determined (without coherent organization, not conceptually coordinated in a single way).
- Even when behavior is apparently well-directed, as when the students quickly respond to each other in constructing worksheet responses, one cannot assign unique justifications—their behavior is a blend of conceptual and physical constraints, not a reasoned plan or articulatable units (breakable into parts).

More could be said along this vein, but consider the implication so far: The theoretical basis of intelligent tutoring systems, namely driving all program behavior by a correct explanatory model of student behavior, is false. No existing model fits the bill, and strong theoretical reasons can be given for the practicality of constructing such a model (on the basis of information available during an interaction) as well as for the theoretical possibility of constructing such a model (on the basis of the non-descriptive aspects of perception and conceptual coordination).

This leads to several possible revisions in the strategy of designing computer-aided instruction systems:

- Incorporate a perceptual categorization model, based on neural networks (e.g., Edelman, 1992), that is, retain the strategy of instruction through explanatory models.
- Tell a different story about the nature and role of cognitive models in instructional systems; for example, say that it's like a teacher with a strong point of view who keeps guiding the student back onto the preferred path (without attempting to exhaustively understand the student's difficulty).
- Reject the use of cognitive modeling for instruction; instead focus on providing a multimedia, hypertext system, perhaps linking student projects through a network.

I can imagine reasonable arguments for each of these alternatives in terms of research goals and practicality. But here I am more interested in the fundamental turning point: The education, psychology, and computer science community cannot proceed with the assumptions that cognitive (student) modeling is useful because it causally explains student behavior. At the level of argumentation, where perceptual details, terminology, objectives, and values are captured in a descriptive language, then there should indeed be a mapping between the program's model and person's representational manipulation—indeed, this is what cognitive modeling has shown. But, and it's a large but, the computer model is replicating the person's expressive behavior (formation and manipulation of descriptive models), not the internal conceptual coordination process. Thus the person's behavior is always more open to blending of perspectives, recognizing of exceptions, and handling of contradictions. And, second, such a set of assumptions does not hold in instructional settings, where perceptual details (where to look), terminology, objectives, and social values are all uncertain and requiring new conceptualization to coordinate.

To restate the conclusion, use of student models in instructional systems should be viewed as being like the use of models in any expert endeavor—a means of classifying a situation so as to conveniently and efficiently determine action plans, without requiring a full understanding of the particulars of a case and their causal relationships. Thus, a librarian can help you find a book

without needing to know the particulars of your motivation. A physician can diagnose and treat a rash without investigating your home environment in detail. And a teacher can provide guidance without understanding how you have gone astray. The point is that expert assistance is always *heuristic*, and need not be scientifically thorough. This relationship between observation, modeling, and action is called *heuristic classification* (Clancey, 1985) and constitutes a description not only of expert systems, but a characterization of human expertise in broad terms (Clancey, 1997b).

Where AI and educational psychologists went astray was to identify the heuristic classification model with human knowledge, suggesting that inference over such models is all that knowledge and reasoning consists of (ignoring the perceptual, cross-model, conceptual coordination aspects). Thus, both the knowledge to be taught and the method for teaching were wrongly identified with descriptive models. A nicely closed system results: The nature of expertise was misconstrued, and consistently, the nature of instructional expertise was misconstrued in the same way. Knowledge consists of more than descriptive models and successful teaching consists of more than manipulating descriptive models (of the student and the domain).

Thus, any of the three alternatives listed above are justifiable: Continue to build an artificial intelligence, if you wish (but you need to understand the nature of conceptualization); tell a different story, viewing the models as "active systems" that guide a student down the well-trodden path (chiefly by being blind and ignorant of alternatives); or find other uses for multimedia technology and models. I like all three alternatives and hope they will each stimulate a broad community of researchers.

Acknowledgments

The Green Globs example presented here was originally presented by Sue Magidson and Judit Moskovitch at a video interaction analysis laboratory at the Institute for Research on Learning during the early 1990s. I produced the transcript and screen shots from the original video. Some of the ideas about dots and lines, including especially the bilingual analysis of "straight" were collaboratively developed in the group discussions at IRL. This work was supported in part by the NASA's Computing, Communications, and Information Technology Program, Intelligent Systems subprogram, Human-Centered Computing element, managed by Mike Shafto.

References

Bamberger, J. (1991). *The mind behind the musical ear*. Cambridge, MA: Harvard University Press.

Bamberger, J., & Schön, D. A. (1991). *Learning as reflective conversation with materials*. London: Sage Publications.

Bartlett, F. C. ([1932] 1977). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Clancey, W. J. (1985). Heuristic classification. *Artificial Intelligence, 27*, 289-350.

Clancey, W. J. (1993a). Situated Cognition: How representations are created and given meaning. In *International Federation for Information Processing, Working Conference "Lessons from Learning"*, Neurope Lab, Archamps, France.

Clancey, W. J. (1993b). Situated action: A neuropsychological interpretation (Response to Vera and Simon). *Cognitive Science, 17*(1), 87-107.

Clancey, W. J. (1997a). *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge, UK: Cambridge University Press.

Clancey, W. J. (1997b). The conceptual nature of knowledge, situations, and activity. In P. Feltovich, R. Hoffman & K. Ford (Eds.), *Expertise in context* (pp. 247–291). Menlo Park, CA: The AAAI Press.

Clancey, W. J. (1999). *Conceptual coordination: How the mind orders experience in time*. Mahwah, NJ: Erlbaum Associates.

Dewey, J. ([1896] 1981). The reflex arc concept in psychology. *Psychological Review, III* (July), 357-370. (Reprinted in J. J. McDermott (Ed.), *The Philosophy of John Dewey*, Chicago: University of Chicago Press, pp. 136-148.)

Dewey, J. ([1902] 1981). *The Child and the Curriculum*. Chicago: University of Chicago Press. (Reprinted in J. J. McDermott (Ed.), *The Philosophy of John Dewey*, Chicago: University of Chicago Press, pp. 511-523.)

Edelman, G. (1992). *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York: Basic Books.

Garrod, S., & Anderson, J. R. (1987). Saying what you mean in dialogue: A study in the conceptual and semantic co-ordination. *Cognition, 27*, 181-218.

Langer, S. ([1942]1958). *Philosophy in a New Key: A study in the symbolism of reason, rite, and art*. New York: Mentor Books.

Lave, J. (1988). *Cognition in practice*. Cambridge, UK: Cambridge University Press.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*(1), 65-100.

Leont'ev A. N. (1979). The problem of activity in psychology. In Wertsch, J. V. (editor), *The concept of activity in soviet psychology* (pp. 37-71). Armonk, NY: M. E. Sharpe.

Norman, D. A. (1982). *Learning and memory*. New York: W. H. Freeman.

Roberts, L. D. (1986). The figure-ground model for explanation of the determination of indexical reference. *Synthese, 68*, 441-486.

Roberts, L. D. (1993). *How reference works*. Albany: State University of New York Press.

Siklóssy, L. (1972). Natural language learning by computer. In H. A. Simon & L. Siklóssy (Eds.), *Representation and Meaning: Experiments with Information Processing Systems* (pp. 288-328). Englewood Cliffs, NJ: Prentice-Hall.

Vera, A., & Simon, H. (1993). Response to comments by Clancey on "Situated Action: A symbolic interpretation." *Cognitive Science, 17*(1), 117-133.