Imagination in Action

Aug Nishizaka
Meiji Gakuin University

Abstract. The psychology of mental imagery has been caught up in the misconception that mental images are entities in an individual. First, I address this longstanding misconception by examining the basic conception of Stephen M. Kosslyn and his colleagues, which underlies all their arguments. This misconception is genuinely conceptual in the sense that it stems from grammatical violations of the use of the words 'see', 'mental image', 'imagination', and so on. Next, in the main body of this article, I attempt to elucidate the concept of imagination by describing practices for organizing imagination. Imagination is re-specified as an organizational property of the ongoing activity rather than any process, event or state in an individual, through the detailed analysis of an occasion in which three 12-year-old participants jointly play a kind of computer game. Finally, Zeno Pylyshyn's 'tacit knowledge' explanation of mental imagery is examined.

Key Words: activity, imagination, interaction, knowledge, mental images

We very often have a mental image. This is a trivial fact. For example, I imagine my father's face, which looks very happy; I imagine a scene in which I have trouble in answering questions after a presentation in an academic meeting, which may be so vivid that I am actually frightened because of the imagined scene. This unproblematic fact, however, may turn into a perplexing one once one asks the following type of question: what *happens* in the mind or in the brain when one has a mental image, or what kind of *state* or *event* is involved when one has a mental image? Certainly, it must be true that a certain (neurophysiological or whatever) state, event or occurrence is involved, and without those (neurophysiological) conditions one could not have any mental image. Problems are caused, however, if one considers mental images to be something happening in some state of the mind or the brain. Mental (visual, auditory or whatever) images are not occurrences or states in an individual. A confusion, in fact, lies in the assumption that when one has an image, the image is something to be *seen*, *heard*, *examined*, *listened to*, and so on, in one's mind. P.M.S. Hacker

Theory & Psychology Copyright © 2003 Sage Publications. Vol. 13(2): 177–207
[0959-3543(200304)13:2;177–207;032183]

(1992), discussing Wittgenstein's ideas on mental images and imagination, invites us to consider the absurdity resulting from this assumption:

> When a composer tells the musicians at a rehearsal that he had imagined *that* chord being played louder, that is not because it *sounded louder* in his imagination, any more than if I expected an explosion to be louder or thought that it would be louder, there was something louder in my expectation or thought. Similarly, when a choreographer tells his dancers that he had imagined a particular *pas de deux* much faster than they danced it, that does not mean that they danced more quickly in his imagination. (pp. 412–413)

In what follows, first, I will address a longstanding confusion in which the psychology of mental imagery has been caught up. This confusion is genuinely conceptual in the sense that it stems from grammatical violations of the use of the words 'see', 'scan', 'mental image', 'imagination', and so on. It cannot be resolved empirically through experiments and observations, precisely because every empirically observable fact is organized by reference to concepts or the use of words. The appropriate use (attribution or avowal) of the words 'see', 'mental image', and so on, is constitutive of the very fact that one sees something, the very fact that one has a mental image, and so on. Next, I will attempt to re-conceptualize imagination and re-specify it as an organizational property of the ongoing activity. At least imaging and imagining on some occasions are best conceptualized as lodged in activity, rather than under an individual's skin. Imagination and what one imagines on these occasions are, far from being concealed from others, transparent and available to others as an interactional resource for the organization of their activity. Imagination as such is organized in a way relevant to the ongoing activity. I attempt to elucidate, through describing practices for organizing imagination, how it is that imagination *is* imagination on a given occasion, rather than specifying any *general* conditions facilitating it.[1]

'Seeing' Mental Images

S.M. Kosslyn performed a series of 'image-scanning' experiments with his colleagues. Kosslyn, T.M. Ball and B.J. Reiser (1978), for example, performed a series of well-known experiments to demonstrate:

> . . . that images are quasi-pictorial entities that can in fact be processed and are not merely epiphenomenal. One of the defining properties of such a representation is that metric distances are embodied in the same way as in a percept of a picture. (p. 53)

In fact, they found that 'more time is required to scan further distances across an image' (p. 50). They take this finding as evidence that mental

images can be processed in the same way as physical pictures. That is to say, the very same thing happens to 'scanning' a mental image as to scanning a picture before one's eyes: it takes more time for further distances. Those results were strong enough for the authors to protect their claim against the possible argument that:

> . . . the apparent effects of distance actually were a consequence of how people accessed some sort of underlying 'list structure'. Parts separated by greater distances on the image might simply be separated by more entries in a list of parts of the object. (p. 47)

One of Kosslyn's earlier experiments uses an image of a boat that has some nameable parts like 'cabin', 'screw', 'bow', and so on. If it takes more time to 'scan' this imaged boat from the screw to the bow than from screw to cabin, it might be due to the number of items (three—screw–cabin–bow—or two—screw–cabin) one has to go through, rather than to the different distances between each pair. The present experiment was designed to eliminate this possibility.[2]

The present experiment:

> . . . involves scanning between the 21 possible pairs of seven locations on an imaged map. Each of these distances was different, and the task seemed sufficiently complex to thwart any attempts to produce intentionally a linear relationship between distance and reaction time. (Kosslyn et al., 1978, p. 51)

Kosslyn et al. (1978) include a detailed description of their procedure. 'The subjects first were asked to learn the locations of the objects on the map [presented to them] by drawing their relative positions.' After learning to draw the locations on a blank sheet of paper without consulting the original map:

> . . . subjects were told that they would hear the name of an object on the map. They were to picture mentally the entire map and then to focus on the object named. Subjects were told that 5 sec after focusing on the named object, another word would be presented; if this word named an object depicted on the map, the subjects were to scan to it and depress one button when they arrived at the dot centered on it. The scanning was to be accomplished by imaging a little black speck zipping in the shortest straight line from the first object to the second. The speck was to move as quickly as possible, while still remaining visible, [and so on]. (pp. 51–52)

There is nothing mysterious about the finding that more time is required to respond to further distances on an image. This finding is, certainly, a 'discovery' in that nothing a priori prevents us from assuming that further distances in the imagination may not require more time. There must be some alternative explanations for this 'discovery'. What I will attempt to do in what follows is not to decide which explanation is valid or more probable. Rather, I argue that the very supposition that underlies Kosslyn et al.'s

(1978) experiments, that is, that mental images are *entities* to be processed somewhere, is *conceptually* confused. As suggested previously, the confusions involved here cannot be resolved empirically by any appropriately designed experiments. They are confusions particularly about the grammar of concepts related to the verb 'see'.[3] Let us begin with Kosslyn et al.'s (1978) claim that 'mental images are quasi-pictorial entities'.

Kosslyn et al. (1978) make an erroneous move when they describe their discovery by saying 'more time is required to *scan* further distances across an image' (p. 50, italics added). We see a picture on the wall, scan distances among objects painted on the picture, and inspect the combination of colors in the picture. However, we cannot see, scan or inspect a mental image in the same way. This 'cannot' refers not simply to a difficulty in doing so (because of special obscurity, instability, etc., of mental images, for example) but to its logical impossibility.

The question to be asked here is not whether or not participants in the experiments actually move a speck on a mental image, but rather whether it makes any sense to talk of moving a speck on a mental image; whether it makes sense to say that I first *see* two points on a mental image and then *scan* the distance between these points. Indeed, it would make no sense to say that one moves something, say an imagined 'black speck', to *inspect* one's own mental image and to *find* something one did not notice, that is, that one *scans* one's own mental image. For imagining something is imagining what one means to imagine, and therefore one *cannot fail to find*, *notice*, *discover*, and so on, something on one's mental image any more than one can fail to know something in what one means. However, this does not mean that we always succeed in finding, noticing, and so on, things on our mental image. I can neither fail nor succeed in finding, noticing, and so on, a stain on my imagined handkerchief. If someone who imagines a handkerchief says, 'I just noticed a stain on the imagined handkerchief after carefully examining it', he or she does not know the meaning of the words 'examine', 'imagine' (and 'image') and 'notice'. P.M.S. Hacker (1992) makes this point:

> I do not doubt that it is King's [College] I imagine on fire, but not because I *recognize* it as an image of King's, have grounds for believing that it is King's I imagine. Rather, I say 'it is King's I imagined', and *this* makes the connection . . . as I say whom I meant or to whom I am writing or what I intend. How do I know? Knowledge is not in question here, nor is ignorance. (pp. 419–420)

Now we can see also a grammatical difference between 'mental image' and 'picture'. I can *compare* a picture (portrait) with what it is a picture of and can tell how the former resembles the latter. On the other hand,

> 'The room is quite different from how I imagined (visualized) it', I may say, and I can tell you how—but not by comparing the room I see with my

> visual image of it (as I might compare the room with a picture of the room). (Hacker, 1992, p. 410)

When I place a picture and its model side by side and compare them, I may discover in the picture a new thing I never noticed. However, I *cannot* discover in my own mental images such a new thing as never noticed by comparing the room I see now before my eyes with my mental (visual) image of it.

Kosslyn and his colleagues appeal to our intuition by citing the following example. When I try to remember the shape of a horse's tail, I may first imagine its whole body, even if roughly, and then *scan* it to the tail. We know now that it is grammatically impossible to *scan* an imagined horse. We must be careful about the grammatical relation between 'remembering' and 'imaging' (or 'imagining'). I do not scan or inspect my image of a horse to remember the shape of the tail. Rather, I just try to remember it by imaging the whole body of the horse, or I just image the whole horse to remember the shape of its tail. Nothing more or less than this.[4]

Some might want to say that this 1970s argument of Kosslyn's is already outdated. However, it seems, the basic confusion about mental images still continues. Indeed, the debate as to whether mental images are 'quasi-pictorial' or 'propositional' (i.e. the so-called 'Imagery Debate') casts its shadow over more recent discussions on mental imagery with its focus shifted. For example, the debate as to whether mental images are ambiguous or unambiguous, which was intensively held from the 1980s through the 1990s, addresses the issue of the difference between seeing physical pictures and 'seeing' mental images.[5]Agreed that mental images are 'neither pictures nor propositions but something in between', they discuss how easily alternative interpretations can be 'discovered' from mental images, or how ambiguous mental images are. The whole issue is, again, to what extent mental images are different from physical pictures. One can easily discover alternative interpretations from physical pictures. Daniel Reisberg, one of the key figures, acknowledging that mental images 'serve as carriers of meaning', most strongly stresses the difference between mental images and pictures and argues for the non-ambiguity of the former. Strikingly, however, even he starts his argument in one of his articles by saying that 'chronometric studies show that the time needed to *scan* across, or to *rotate*, or to *zoom in* on a mental image all correspond to what one would expect with a pictorial representation' and that these 'findings, showing that images preserve the metric properties of space, speak directly to the claim that images are indeed depictive' (Reisberg, 1996, p.120, italics added).[6] He also speaks of 'discovery from imagery', attention to an image ('attended parts of image'), the density and size of an image, and so on. Just as one simply imagines that a black speck moves on a figure rather than scanning across the image of the figure, so one simply imagines that a figure rotates rather

than rotating the image of the figure, one simply discovers an alternative construal of a *figure* by trying to remember the figure with imaging rather than discovering another construal of a *mental image* (of the figure) by inspecting the image, and so on. The problem here is caused by confusions about the object of imagining or imaging. The object of imagining or even imaging is not a mental image but the figure whose image the mental image is.[7] One cannot pay attention to parts of an image but only to parts of the figure by imaging it. One cannot talk of the size of an image; one can only imagine, for instance, that one sees a horse close to it or far from it or that a figure is drawn larger than another.[8]

In what follows, I will focus on imagination, rather than mental images as such. Imagination tends to be regarded as having a mental image, but as Alan R. White (1990) repeatedly emphasizes, having a mental image is not essential to imagination. Indeed, one often imagines the 'non-imageable and non-picturable'; one can imagine, for example, that the Crucifixion is the redemption of human beings, or how to solve a mathematical problem. Furthermore, even when one imagines something by having its image, having the image is not imagination per se. Imagining Mary is different from imagining her identical twin Martha even though their images are identical. It is crucial to avoid conflating imagining and having a mental image. Their conflation causes further serious confusions.

Imagination involves, in one way or another, what is absent right here and now (sometimes we can even imagine what does not exist in this real world, though, of course, we often imagine what is real). Where, then, is the imagined thing that is absent but present to me right now as well? It must be in my mind or in my brain—this is a common fallacy. If you imagine that Peter who is now in Berlin is in Paris, the object of your imagination is now in Berlin, not in your head (cf. Sartre, 1948). If you imagine that a centaur runs on Santa Monica Boulevard, the object of your imagination is nowhere. I suspect that the fallacy originates from that conflation of imagination and having a mental image. We are inclined to consider mental images as hanging somewhere in the mind or in the brain like pictures hanging on the wall, and consider what one imagines in one's mind as 'hanging' somewhere in one's mind. However, images, and imaginative contents (e.g. that Peter is now in Paris) in general, do not exist in such a way that one can inquire about their location. It does not make sense to speak of their location. They are *neither* somewhere *nor* nowhere.[9]

What one imagines and, as will be seen later, imagination as such, too, are not located in the mind or in the brain. In what follows, I attempt to demonstrate that, and how what one imagines and imagination on an occasion are rather an organizational property of the ongoing activity. They are organized in a way appropriate to each current phase of the ongoing activity, that is, in the way that it is available to others as an interactional resource for the organization of the activity at each moment. What one

imagines and how one imagines it are not concealed from others there. I suggest that the expression 'imagination' is not the label of any esoteric state or occurrence within an individual.

A note is in order here about the logical status of the demonstration in the following section. I attempt a detailed analysis of a scene from videotaped interaction. The fact that I analyze empirical material does not mean that I attempt to *prove* a hypothesis or whatever. I do not attempt to discover 'new' things, which we do not know. The material is intended as a 'reminder' of *what we already know* about imagination or practices we are already well acquainted with for accomplishing imagination.[10] The analysis is an attempt to elucidate our concept of imagination and explicate our practices for accomplishing imagination. Its purpose is to obtain a 'perspicuous representation' (Wittgenstein, 1953) of what we know about imagination.

It may not necessarily be obvious how explicating practices and elucidating concepts are connected to each other. The concept of imagination is provided by the *use* of the expressions 'imagining', 'imagination', and so on. The use of expressions is always sensitive to all the particulars of each situation. Since this is the case, the detailed analysis of how people actually do use the expression in question and other related expressions in each unique situation should be the best reminder of the concept in question. Explicating practices is thus connected to the elucidation of concepts. Certainly, in the segment to be analyzed in the following section, the participants do not explicitly use the word 'imagine', 'imagination', or the like. However, they are in a situation describable as one in which they *imagine* something, that is, they are in a situation so describable not only for analysts but also for the participants themselves. The focus of the analysis is on how the participants accomplish this describability by employing various practices within the very situation so describable (see note 13).[11]

Imagining within Activity

The case I want to examine clearly illustrates the point being made here. It is taken from a multiple-participant activity. In jointly performing an activity, participants often have to know what each other imagines and they often have to imagine what is relevant to the ongoing activity in a relevant way. They use each other's imagination to collaboratively organize their activity. Indeed, what participants imagine in the following fragments is before their eyes, rather than before their 'mind's eye'.

The fragments are taken from an audio-visual recording of three teenagers jointly playing a 'computer game', called 'AlgoBlock'.[12] Players 'write' programs to move a submarine on the computer screen, by connecting a number of approximately 10 cm × 10 cm × 10 cm cubic blocks to each other on a table in an order they choose to accomplish the desired submarine

movement (a command like GO-FORWARD, GO-RIGHT, GO-LEFT, ROTATE, or the like, is allocated to each block). Once they have done this, they press the 'play' button on the controller connecting the blocks to the computer to see if their instructions actually produce the turns and other desired movements of the submarine on the computer screen.

The designers of the system provided the videotape. They performed an 'experiment', and videotaped it with two cameras. In the experiment, three 12-year-old participants were assigned three tasks and in each task asked to move the submarine towards the goal. Two instructors attended the experiment to switch tasks and to assist the participants in the events of system trouble. Figure 1 reproduces the layout of participants, cameras and other tools, which can be seen on the video. The command blocks, the controller, a personal computer and a large-size monitor are placed on one table. The participants surround the table on which the blocks lie. In what follows I call the participants A, B and C as shown in Figure 1.

The fragments to be examined here are excerpted from a scene less than 30 seconds long. I focus on how the participants collaboratively complete the current task. The original purposes of the system and the experiment the designers of the system performed are not relevant here. In the fragments the players discuss how to set up the angle of a turn at a point on the computer screen to send the submarine to its goal. In these fragments, the computer screen displays the submarine remaining stationary at the home position.

As is evident from the above description,[13] there are two distinct operational fields relevant to the ongoing activity, that is, the computer screen and the blocks on the table. They are salient not only for external observers but for the participants themselves as well. Indeed, the participants often point with their index fingers and gaze at the screen and the blocks. What is important, however, is how they *publicly* (i.e. in a way accessible to each other) manage their orientations to those two operational fields with their various bodily movements, including pointing gestures and gaze

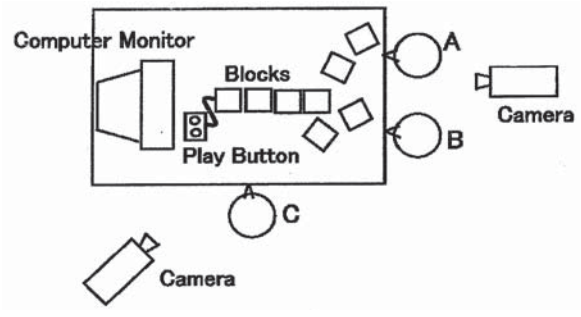


FIGURE 1. The layout of participants, cameras and other tools.

directions, in order to *jointly* accomplish their ongoing activity. For their public management of their orientations, each other's body constitutes another unmarked operational field, which is also available to each participant as an interactional resource for accomplishing their joint activity, that is, jointly 'writing a program' for moving a submarine on the computer screen.

In the following fragments, the participants often stretch their hands and slide them along the computer screen. As is obvious on video, their hand movements draw imagined routes a submarine might take on the screen. The question should be how this very fact (that hand movements draw imagined routes of a submarine) is achieved. This fact is not automatically provided by 'mere' hand movements (even with the mover's intention). It is rather a joint accomplishment by participants. In what follows, first I describe experiential properties of those imagined routes of the submarine, and refine the question that should be asked. Then, through the detailed analysis of those fragments, I demonstrate how the participants jointly achieve their imagination in their ongoing activity, by examining some of their practices of managing orientations within interaction.

Hand Movements and the Structure of the Environment

As mentioned previously, the participants discuss how to set up the angle of a turn of the submarine. They, indeed, argue against each other. In Fragment 1, all the participants mark their utterance as a counter-argument with the challenge token '*datte* [why]' placed at the beginning of their utterance:[14]

Fragment 1 [ll. 1–4a; The Original Japanese Transcript and Its Free Translation]

1	A:	*Datte*	*asoko*,	*koo*	[*it*	*te,*	*koo-*	
		why	there	this way	*go*	*and*		
2	B:				[*Datte*	*soo*	*shi*	*tara*
					why	so	do	if
3	?:	[( )						
4	C:	[*Datte*	*nana*	*ko*	*chokushin*	*shi*	*te*	
		why	seven	piece	going straight	do	and	
4a		*kyuujuu*	*do*	*de*	*magaru*	*deshoo?*	*Itsutsu.*	
		ninety	degree	at	turn	P	five pieces	

1 A: Why, there ((it)) goes this way //and this-
2 B: Why, if so
3 ?: //( )
4 C: Why, ((it)) goes seven pieces straight, and turns at
4a ninety degrees, right? Five pieces.

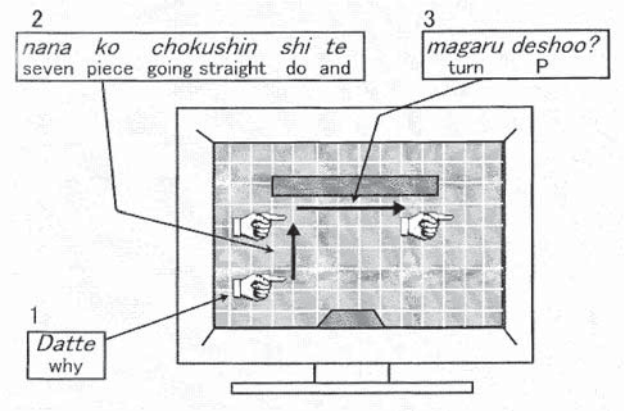


FIGURE 2. Line 4 in Fragment 1.

In arguing against each other, they support their own claims with their imagined routes of the submarine. For example, C at line 4, while saying '*datte* [why]', points her left index finger at the screen and, precisely when she says '*nana ko chokushin shi te* [((it)) goes seven pieces straight]', she slides her left hand upward along the screen. Then, saying '*magaru deshoo?* [((it)) turns, right?]', she slides the same hand to the right (see Figure 2). Insofar as it is of the submarine that she says 'it goes straight' and 'it turns' and the submarine *actually* remains stationary at its home position, C's hand movements draw an *imagined* route the submarine might take. In other words, C, with her hands, that is, in a way visible to others, imagines possible routes the submarine would take in accordance with her proposed layouts of blocks. C demonstrates what route the submarine would take using imagination in front of others (rather than in her head!).

Let us turn to line 1 to further describe the phenomenological properties of imagined submarine routes. A at line 1, when she says '*asoko* [there]', looking at the computer screen, stretches her right arm and points her index finger at the screen. Then, when she says '*koo* [this way]', she slides that finger along the screen. In pointing her index finger at the screen in this way, she enhances her orientation to the screen in a way noticeable to other participants (as can be seen from Figure 1, participants position themselves such that they can see where each other is oriented).[15] It should be noted, however, that A's utterance '*koo* [this way]' does not refer to something on the screen. Rather, it seems to refer to the hand movements as such. Her use of these hand movements seems to be the same as the one made of a hand gesture, for example, to show the size of something saying, 'I'm looking for a purse of *this* size.' However, to say this is not enough to capture the properties of the hand gesture. It is already obvious to us, who know that C's

hand movements at line 4 draw an imagined route of the submarine, that A's hand movements also draw an imagined route of the submarine. Certainly, the participants see imagined routes precisely where hands are moving. However, those hand movements could not constitute an imagined route in isolation. When A stretches her hand to the screen saying '*asoko* [there]', the computer screen comes into the others' view as well as her hand itself (note that '*asoko*' evidently refers to some part of the screen). The screen is a field for the submarine's movements, and it has a unique structure; some obstacles for the submarine are placed on it and even a grid (to assist the players in 'programming') is drawn on it. It is only with *both* the gestures *and* the structure of the screen in view that the participants (and we as analysts) can see an imagined route just where the hand is moving. Charles Goodwin (in press) calls these 'symbiotic gestures'. According to Goodwin, 'symbiotic gestures' are organized through the conjunction of talk, gestures and the structure of the environment. These things elaborate each other as a unique 'semiotic field' through highlighting some specific aspects of each other to produce a specific object such an imagined route of the submarine (see also Goodwin, 2002a, 2002b, among others).

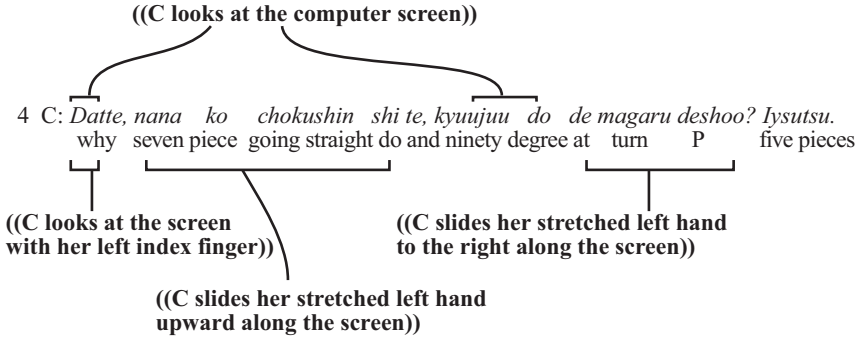
Now some points need to be clarified here. First, certainly, the participants' hand movements constitute imagined routes of the submarine. But this fact is not so simple.[16] It is because the submarine on the screen in front of them does not move that those hand movements draw its *imagined* routes, and, on the other hand, these are still imagined *routes of the submarine*, that is, it must be nothing but submarine routes that they imagine with their hands. That is to say, on the one hand, those possible submarine routes are not only now absent but also are 'posited as absent' (cf. Sartre, 1948). On the other hand, generally speaking, there are indefinitely many things that are now absent. The question to be asked here is how this specific object is posted as specifically absent on this specific occasion. Second, we saw that imagined routes are the product of the organization of hand gestures as 'symbiotic gestures', that is, that they emerge through the conjunction of talk, gestures and the structure of the environment. Then the above question can be formulated thus: how is this conjunction organized in the way specifically appropriate to this particular occasion? Finally, as suggested previously, this conjunction is built in a specifically appropriate way through the management of the participants' orientations to those aspects of the structure of the environment that are specifically relevant to each current phase of the ongoing activity. Particularly, those two operational fields relevant to the ongoing activity, that is, the command blocks on the table and the computer screen, are located such that the participants cannot look at them simultaneously (see Figure 1). If both those hand gestures and the computer screen are in the participants' view but they are not related to the blocks as well, those gestures would never produce that specific object, that is, an imagined (now absent) route. Now the question should be how the

participants *publicly* (i.e. in a way accessible to each other) coordinate their orientations to those operational fields relevant to the ongoing activity in a specifically appropriate way in the real-time development of *interaction*, such that their hand movements are organized as imagined routes of the submarine.

Coordinating Orientations within Bodily Conduct

Imagined submarine routes embody the 'meaning' of possible layouts of the blocks on the table. That is to say, to produce imagined routes of the submarine, the participants have to relate, in a publicly visible way, those two operational fields relevant to their current activity, that is, the computer screen around which the submarine should move when they press the 'play' button and those blocks with which they 'write' programs to move the submarine on the screen. However, as I said above, those two operational fields are spread out in different directions from each participant, such that they have to orient themselves differently when monitoring the screen than when handling the blocks. Particularly, C has to radically change her facial orientation to turn to the computer monitor from the blocks on the table. However, she takes advantage of this position provided to her by the layout of the setting to relate those two operational fields in front of other participants. C at line 4 in Fragment 1 basically looks at the blocks, and only looks at the computer screen momentarily with the front of her trunk fixed toward the blocks, while keeping her index finger on the monitor screen. Fragment 1a shows the relation between C's (original Japanese) utterance and her hand movements at line 4 (see also Figure 3):[17]

Fragment 1a [l. 4; the Original and a Phrase-by-Phrase Translation]



The orientation displayed by her outstretched hand and the orientation displayed by her face (and by her upper body) are integrated in her body.

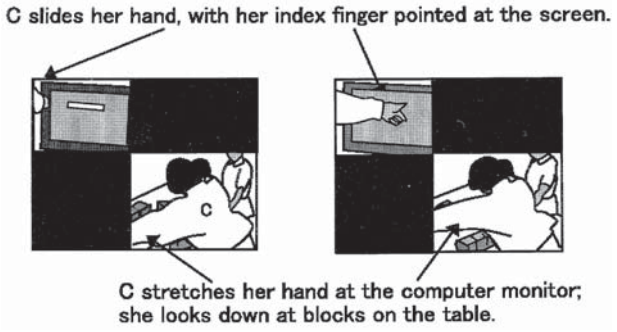


FIGURE 3. Line 4 in Fragment 1.

Those two visual fields, that is, the monitor screen and the blocks, are publicly connected to each other through her gestures and postures.

Note that C at line 4 refers to a possible distance the submarine would go with '*nana ko* [seven pieces]'. The Japanese word '*ko*', which is a suffix to numerals, is hardly used to describe distances but rather used to count solid discrete things. (In Japanese we append different suffixes to numerals for different types of things. For birds, we would say, '*nana wa*', for people, '*nana nin*', for books, '*nana satsu*', for instance. You might be reminded that English speakers would say 'seven slices of bread', 'seven sheets of paper', and so on.) However, insofar as the expression 'goes straight' still describes the submarine's possible movement, C appears to refer to the number of the estimated *steps* the submarine would take, using the suffix '*ko* [piece]'. More precisely, she appears to refer to the number of those squares of the grid on the screen that measure a possible distance of the submarine's movement. That is to say, she actually refers to the number that is relevant to operations on the blocks, that is, the number that the participant may set up on a block. This measurement of a possible distance of the submarine's movement emerges as the product of the juxtaposition of talk ('*ko* [piece]'), the hand gestures and the structure of the screen (a grid and obstacles on it), all then referring back to a possible operation on the block. Thus, an imagined route of the submarine is produced through the mutual highlighting of some specific aspects of talk, hand gestures, posture, the blocks and the structure of the screen.

Coordination between Bodily Conduct and Talk

Let us now turn to another fragment and later return to line 1 in Fragment 1:

Fragment 2 [ll. 12–16; The Original Japanese Transcript and Its Free Translation]

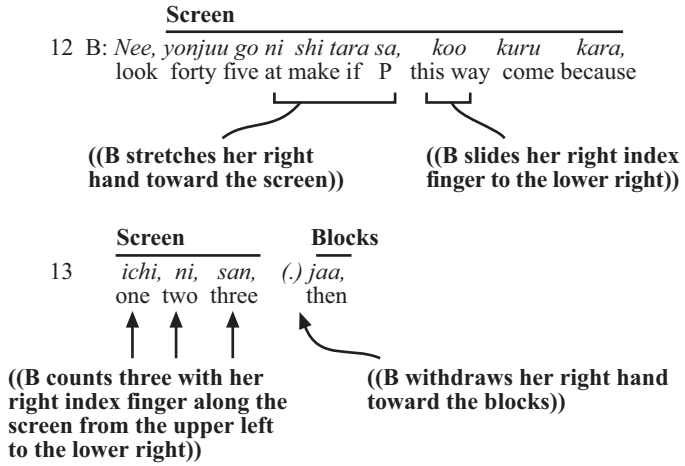
12	B:	*Nee*	*yonjuu*	*go*	*ni*	*shi*	*tara*	*sa*	*koo*	*kuru*	*kara*
		look	forty	five	at	make	if	P	this way	come	because
13		[*ichi,*	(.) *ni*,	[(.) *san*,	(.) *jaa*						
		one	two	three	then						
14	C:	[*Demo*	sa,		[*kore-*						
		but	P		this						
15	A:	*Mittsu*	*ni*	*shi*	[*tara::*						
		three pieces	at	make	if						
16	C:	[*De*	*sa*,	*kore*	*wa*	*sa*,	(.) *nana*	*ko*	*it*	*te*	
		and	P	this	P	P	seven	piece	go	and	
16a		*kyuujuu*	*do*	*ni*	*kaete*,	*chokushin*	*ni*	*shi*	*te::*		
		ninety	degree	to	change	going straight	P	make	and		

12 B: Look, if ((you)) make ((it)) forty-five, ((it)) comes this way,
13 //One, two, //three, (.) then,
14 C: But,
14 C: this-
15 A: How about three pieces,
16 C: And, this is, (.) ((it)) goes seven pieces, and change ((it))
16a to ninety degrees and make ((it)) straight

First, B's conduct at line 12 is prominent in one respect. It is evident from the context that 'forty-five' signifies a possible angle of the submarine's turn which should be set up on the blocks. Indeed, the phrase '*ni shi tara* [if ((you)) make ((it))]' in her utterance '*yonjuu go ni shi tara* (if ((you)) make ((it)) forty-five]' explicitly refers to an operation on the blocks. When saying this, she stretches out her right hand toward the computer screen. That is to say, B displays her orientation to the blocks with her utterance, and at the same time displays her orientation to the screen with her hand movement. Then, immediately after this, she slides her stretched hand along the screen, saying, '((it)) comes this way'. The fact that this hand movement draws an imagined route of the submarine in the same way as A's and C's hand movements in Fragment 1 is, again, accomplished through orientations to those two operational fields relevant to the participants' ongoing activity being publicly connected to each other. In this case, the coordination of

orientations is achieved through the sequential context, that is, through conducting the hand gesture immediately after establishing the relation of orientations to the two operational fields, as shown in Fragment 2a.

Fragment 2a [ll. 12–13; The Original and a Phrase-by-Phrase Translation]
(Solid lines on B's utterance indicate her facial orientation.)



Immediately after sliding her right index finger and saying, '*koo* [this way]', B counts three with the same (right index) finger, moved in the same direction as just now but waved this time. In so doing, she takes the distance of the imagined route just drawn and highlights (according to a grid on the screen) *steps* the submarine might take; these *steps* are, as we saw previously, precisely units that are relevant to operations on the blocks. Indeed, immediately after this B, saying, '*jaa* [then]', withdraws her right hand away from the screen and turns to the blocks (see Fragment 2a). That is, she shifts her orientation from the computer screen to the blocks, responding ('then') to her own counting *steps*. Here, those aspects of the screen that are highlighted by hand gestures *plus* counting are sequentially connected to possible operations of blocks. Thus, the meaning of her hand movements (i.e. their organization into an imagined route) is elaborated through the orientations to the computer screen being publicly hooked onto the orientations to operations on the blocks.

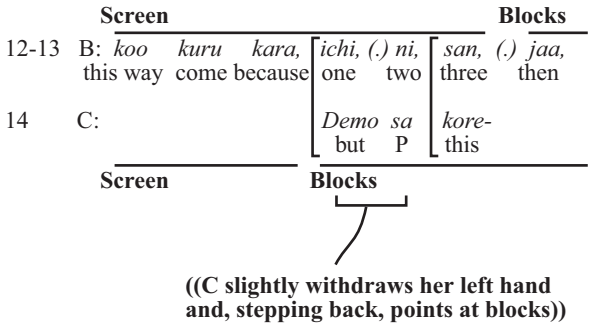
Coordination Distributed between Participants

What has been said about participants' management and coordination of their orientations to the operational fields relevant to their activity focuses on

a single participant's conduct. It should be noted, however, that all the conduct so far examined is interactional in the sense that management and coordination of orientations must be achieved in a way visible and accessible to all the participants jointly engaging in the ongoing activity.

Now I turn to a case where the management and coordination of orientations to the two operational fields is prominently distributed on more than one participant's conduct. At line 14 in Fragment 2, C attempts to argue against B's argument just provided (that they should make the submarine turn at forty-five degrees), with a disagreement marker ('*demo* [but]'), overlapping with B's counting. C's attempt at the counter-argument is marked also by her salient bodily movement (see also Figure 4):

Fragment 2b [ll. 12–14; The Original and a Phrase-by-Phrase Translation]
(Solid lines below C's utterance indicate C's facial orientation.)



When C says, '*demo* [but]', at line 14, she steps back slightly, but markedly, to secure in front of her a *space* for manipulations of the blocks, and, shifting her gaze from the computer screen to the blocks, she points at the blocks with her index finger, with which she pointed at the screen a moment ago (see also Figure 3). That is to say, while B advances an argument in reference to the screen, C is arguing against it in reference to the blocks on the table. There are two things to be noted. First, C makes the counter-character of her counter-argument salient by contrasting her orientation to the blocks with B's enhanced orientation to the screen displayed by her face and hand movements. Second, a counter-argument is generally expected to come only at a second position after an argument, claim or proposal at a first position. C is counter to B's demonstration on the screen in reference to the blocks. Thus, B's orientation to the screen is sequentially connected to C's orientation to the blocks in accordance with the utterance-connection format 'argument–counter-argument'.

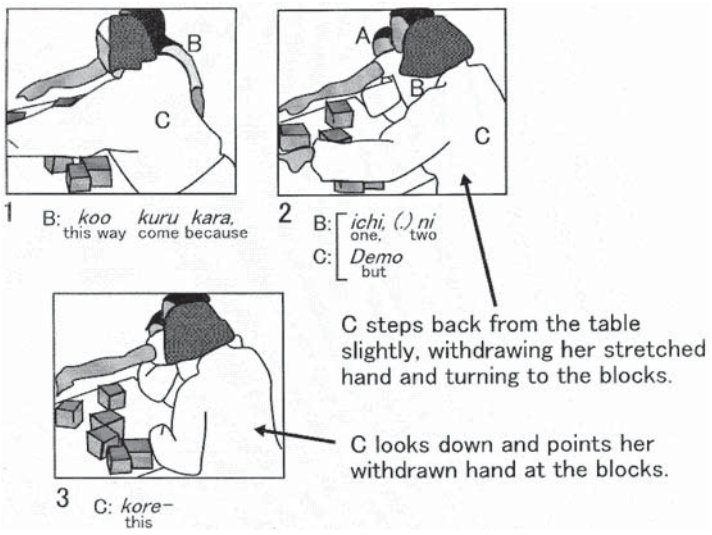


FIGURE 4. Lines 12–14 in Fragment 2.

Now back to A's conduct at line 1 in Fragment 1. Certainly, looking at this conduct of A's in isolation, it is difficult to find any specific practice of coordinating her orientations to those two operational fields. She only slides her stretched hand along the computer screen, uttering a deictic term (i.e. '*koo* [this way]'). However, through the placement of this conduct of hers in the sequential context, her hand movement obtains a clear meaning. As I said previously, it is located in the sequential context of the argument. That is to say, it is designed to be a challenge to another participant, as marked by the challenge token '*datte* [why]'. In fact, the preceding argument by B (not reproduced in the transcripts; here, too, B argues that they should make the submarine turn at forty-five degrees), to which C asked a follow-up (perhaps challenging) question to which A challenges at line 1, was accompanied by the same hand gesture (sliding her hand along the screen). Moreover, this hand gesture (of B's) is organized into an imagined route through the very same practice as the one at line 12 that we saw above. In this sequential context, A's hand movements at line 1 is understandable (to the participants themselves and analysts) as drawing an imagined route of the submarine. Such understanding by the participants is shown in the subsequent conduct by C at line 4. On the one hand, as we have seen, when here at line 4 C defended herself against other participants, her hand movements are organized as drawing an imagined route through her talk, hand movements, posture and face orientations (see Fragment 1a and Figure 3). On the other hand, when C defends herself, she appears to pick up A's hand gestures, that

is, pointing at the screen and then sliding her hand along it (in fact C's hand movements start immediately after A's end) and try to ground her self-defense in the imagined route being produced in the way I previously described. Thus, A's hand movements at line 1 are also organized into an imagined submarine route through the placement of her conduct in the interactional format 'challenge–defense/counter-challenge'.

We have so far seen that imagined routes, specifically posited as absent, are produced through the conjunction of talk, hand gestures and the structure of the screen, in which these elements elaborate each other and highlight each other's specific aspects, and that such conjunction is achieved through management and coordination, within a real-time development of the ongoing activity, of the participants' orientations to those operational fields relevant to the activity.

Discussion

Seeing What One Imagines

Imagined routes of the submarine are not mental images. The players of 'AlgoBlock' *see* those imagined routes of the submarine in front of them with their eyes (not 'see' them with their mind's eye). It is not that what they actually see is only their hand movements and that the 'meaning' of a possible route of the submarine comes after this, perhaps, through interpretation. Rather, the imagined routes *are* what they see. There is nothing mysterious about this. For example, airlines may provide a map on which they draw red lines to show their routes. These red lines are imagined air routes (insofar as airplanes do not actually move along the lines). Here, too, one simply sees imagined air routes, not just red lines (not that one first sees just red lines and then interprets them as imagined air routes). In these terms, those imagined routes, drawn with hand movements or red ink, have a grammatically different relation to seeing than mental images one has in one's mind.[18] However, it should be noted that there is a salient feature of imagining something with a pen or with paint, which has to do with the grammar of 'imagination'. I draw a picture of an imagined animal with a pencil. I now see the picture before my eyes. I notice that the eyes of the animal are not placed symmetrically. Looking closely, I *infer* that it has thirty-two teeth, including occluded ones. I *discover* that the lines around its neck are distorted because of a crease in the sheet on which I drew it. But I cannot notice, guess or discover that the animal has a horn, two eyes, teeth, and so on: not because it does not have those things (surely it has!), but because it does not make any sense to say, 'I did not notice it has a horn before looking closely.' When I see a picture of the Crucifixion, I may notice or find that there are a lot of scars due to flogging all over Jesus' body.

Would it have been possible, however, for the painter, Grüenewald, himself to have noticed or found, right after he had finished his painting, that there were a lot of scars on Jesus' body? Imagination is always imagination *of something*. When a participant in 'AlgoBlock' imagines with her hand that the submarine would go from the left to the right, she cannot notice or find that the imagined route goes from the left to the right by looking at that imagined route she is now drawing. It may be because of this 'intentionality' of imagination that one cannot notice, find, infer, and so on, something in mental images.

The Activity of Scanning

Participants in 'AlgoBlock' show imagined routes to each other and, in so doing, they collaboratively organize their activity. Their hand gestures, which draw imagined routes, are also organized within the ongoing activity into an activity of *scanning* the computer screen (a material image). By sliding their hands along the screen, the participants try to find at what angle the submarine would take a turn and how many steps it would take in accordance with each proposed 'program', that is, each proposed layout of blocks. For example, in Fragment 2, B scans distances on the computer screen by counting (according to a grid on the screen) the number of steps the submarine might take. She moves her stretched hand from the upper left to the lower right and counts three. In so doing, she highlights in front of other participants those aspects that are relevant to the problem at hand, that is, how to lay the blocks out. Her drawing an imagined submarine route is scanning a possible distance the submarine would take by moving (not a 'black speck', but) her hand.

By drawing imagined routes, the participants not only demonstrate the 'meaning' of their proposed layouts of blocks to others, but also make visible those aspects that may be difficult even for the proposer herself to discern without any auxiliary means. The number of steps to be taken after the proposed turn may be one of those aspects; it may be easier to discern if one moves her hand over the screen, that is, if one actually draws an imagined route. Thus, drawing imagined routes constitutes a crucial resource for making visible those aspects relevant to the current activity in an appropriate way in the actual course of interaction. There is nothing for the participant herself to discover, find, notice, and so on, about what she meant to imagine. When she imagines the route the submarine would take if they set up on the blocks the angle of the turn in question at forty-five degrees, it does not make any sense to speak of her *discovery* that the route goes from the upper left to the lower right, for example. However, if she 'applies' an imagined route to the screen, she may discover the number of steps the submarine would take in an appropriate way to the ongoing activity, such

that she may set up an appropriate number on a block. In fact, at line 12 in Fragment 2, B responds to her own counting by saying 'then [*jaa*]' and turns to the blocks on the table; in so doing, she makes it publicly available that she now goes on to operate on the blocks after doing *finding*, or doing confirming, the number to be set up.

Imagination as an Organizational Property of the Ongoing Activity

Surely, the participants' imagination is organized in an appropriate way to the current activity they jointly engage in, that is, jointly 'writing' a program to move the submarine. The most important thing to bear in mind, however, is that the participants do not need to engage in an imagining act along with their ongoing activity. Indeed, there is no need for any separate act of imagination besides their ongoing activity that facilitates the production of imagined routes. Neither are their hand movements as such acts of imagination. Rather, through those hand movements being in an appropriate way part of the ongoing activity, imagination is achieved, that is, those hand movements are organized into imagined routes. In this sense, imagination is an organizational property of the ongoing activity.

I do not know that this is true for all possible instances of imagination. However, we can have some sense about the possible essential relationship between imagination and activity, which has been neglected in the literature. Even when you read a book imagining the details of a scene described in it, or try to solve a mathematical problem by imagining a variety of possibilities, your imagination is still achieved in a way appropriate to the activity you currently engage in (i.e. reading a book, or trying to solve a problem). In fact, any impressions, feelings or sensations of colors, sounds, faces, numbers, and so on, may cross one's mind at any time. However, the question is when exactly one can be said specifically to *imagine something* and when one can be said specifically to *have a mental image*. It is not every time one has those impressions or feelings. To say specifically that one reads a book imagining how the described scene looks is to say something more than that one reads a book only with some impressions of the scene popping up in one's mind. How about reveries? The question is the same: when can one be said specifically to have a reverie, indulge oneself in imagination, and so on? Probably, it is negatively related to the ongoing activity. When the activity one currently engages in is *specifically* interrupted by something reportable as absent, for example, one can be said specifically (by oneself or others) to have a reverie or indulge oneself in imagination. Thus, even when there is no interaction with others, specifically imagining something may still be an organizational property of the ongoing activity.

An Excursion: A Descriptionalist View

Kosslyn and his colleagues (1978) took their results ('more time is required to scan further distances across an image') to confirm their claim that the representation underlying images is basically depictive, that is, quasi-pictorial. Pylyshyn (1981) provides an alternative explanation, implying that the results of Kosslyn et al.'s experiments cannot sufficiently support their claim.

After drawing our attention to the distinction between (1) focusing on a certain object in a mental image and moving a black speck as quickly as possible to decide whether a second named object is present in that image, and (2) imagining yourself in a certain real situation in which you are focusing on a certain object in a certain scene and imagining that a black speck moves across the scene to a second named object (p. 32), Pylyshyn notes that in the image-scanning experiments:

> . . . what subjects were required to imagine was a real physical event (since terms like *move* and *shift* refer to physical processes) about the duration of which they would clearly have some reasonable tacit knowledge. For example, they would know implicitly that it takes a moving object longer to move through a greater distance. (p. 33)

Then, his alternative explanation follows: 'they will naturally attempt to reproduce a temporal sequence of representations corresponding to the sequence they believe would arise from actually viewing the event of scanning across a scene' (p. 34). If one imagines that a black speck moves on a map, one will often imagine this such that more time is required for the speck to move across further distances, as is actually the case. The 'reason I imagine things happening more or less the way that they actually do is . . . simply because I know how things generally happen—because I have been told, or have induced, what some of the general principles are' (p. 39). Indeed, one may be unaware that one imagines the move of a speck in the way that it would actually move across a real scene, and may be surprised to be told that one imagines it exactly in this way. It is, Pylyshyn argues, because the knowledge of those 'general principles' of how things happen is *tacit*. 'I have a tacit physical theory which is good enough to predict most ordinary everyday natural events correctly most of the time' (p. 39).

My criticism of Kosslyn's *grammatical* confusions might seem close to Pylyshyn's criticism of his *explanation*, insofar as the latter casts doubt on Kosslyn and his colleagues' claim that the participants in their experiments really *scan* and *retrieve information from* mental images. However, Pylyshyn's (1981) argument is also deficient from the grammatical point of view. He thinks that when we have a mental image, 'propositional' (rather than depictive or quasi-pictorial) information is processed. In his earlier article (Pylyshyn, 1973) he says this:

> Just because we *know* that we use certain mnemonic strategies . . . or that we 'see' certain objects in our 'mind's eye' or 'hear' ourselves rehearsing a series of numbers, etc., we cannot assume that the contents of such subjective knowledge can be identified with the kind of information-processing procedures which will go into an explanatory theory. (p. 3)

According to Pylyshyn, the expression 'seeing in one's mind's eye' is a quite ordinary one and one does not have to deny that we 'see' certain objects 'in our mind's eye'; but it does not follow that we have to admit that we actually *see* them in our mind's eye in the literal sense. We do not see them in our mind's eye in the same way as we see a picture on the wall before our eyes. When we 'see in our mind's eye', Pylyshyn argues, a quite different kind of information than the 'depictive' or 'quasi-pictorial' one is actually processed. Now the question to be asked is this: what is the logical status of the kind of information-processing procedures that cannot be identified with 'the contents of such subjective knowledge'?

This type of idea is subject to G.P. Baker and P.M.S. Hacker's (1984) criticism of the notion of tacit knowledge. They note that the fact that 'our speech *can* (perhaps) be mapped on to a complex calculus no more shows that we have been operating one than the mere possibility of mapping Zulu war dances on to chess shows that Zulu warriors are chess-players'. What are 'the grounds for attributing to a speaker such tacit knowledge'? If the speaker's 'correct discourse is *all* that shows his tacit knowledge and incorrect discourse is *all* that shows ignorance, then the hypothesis that he can produce and understand sentences because of such tacit knowledge is both untestable and vacuous' (p. 341). Certainly, we may write a program that produces only correct discourse (and not incorrect discourse), but there is still no ground for attributing to us *this* program as 'tacit knowledge'.

Indeed, the model Pylyshyn advanced in his 1973 article *requires* some knowledge to be 'tacit'. There Pylyshyn argues:

> . . . the fact that as novel instances are presented to him, a subject can keep coming up with clever new heuristic procedures for assigning those instances to concepts . . . suggests that his representation of the concept is not limited to a finite list of such consciously available procedures. Rather, he is able to creatively generate new heuristic procedures from a representation which, while it is most likely procedural, is itself more abstract than a list of the procedures he is aware of using on specific occasions. (p. 17)

First, the idea of 'tacit knowledge' (a procedural representation 'which . . . is itself more abstract than a list of the procedures he is aware of') is motivated by considering our 'creative' ability to conceptualize 'novel instances'. How can we comprehend 'novel' instances that we never encountered before? This problem is basically the same as what is called the 'new sentence problem' in the Chomskian theory of generative grammar. We can produce and understand 'novel' sentences (or an indefinitely large number of

sentences, most of which are 'new'). How is it possible to do this? To solve this problem, rules (or procedures) that subsume an indefinitely large number of sentences (or instances) are required. These rules, which divide *all* the possible sentences into grammatical and ungrammatical sentences, it is said, could make it possible to produce and understand 'novel' sentences. Second, these rules (or procedures) could not be learned by experience, insofar they are beyond a finite set of sentences (or instances) we could ever come across actually. Rather, it is said, they are innate or embedded in the nervous system. That is, we know them only 'tacitly'. They will not be learned or lost.

Many have already pointed out the fallacy involved in this Chomskian argument (see especially Baker & Hacker, 1984; Coulter, 1979, 1991). First, does it make sense to talk of rules 'known' in this way, that is, embedded in the nervous system? To follow a rule in the ordinary sense implies the ability to refer to the rule to explain one's own or other's behavior, to deliberately violate the rule, and so on. It is impossible to say of 'rules' embedded in the nervous system that we can use (or even exploit) them in this way. We cannot *follow* these 'rules'. Second, as Coulter (1991) argues, the 'new sentence problem' is a pseudo-problem. Does it make sense to ask *how it is possible* to produce and understand sentences? The question 'How is it possible?' is usually asked when one is confronted with more or less difficult circumstances. For example, how is it possible to get the university administration to accept our demands? We then work out negotiation tactics, carefully consider alternative plans and put one of them into practice. If one asks, 'How is it possible to produce and understand sentences?', the production and understanding of sentences turn into more or less difficult tasks to be achieved with the making and practicing of plans. Of course, one who is not familiar with physiology may encounter a 'novel' expression like 'The amygdala projects directly to the hypothalamus.' Then trying to understand this novel expression, one may make a plan and put it into practice (one may go to the library and consult some relevant dictionaries, for example). However, we do not do this in order to understand ordinary sentences that are rather *familiar* enough to us, even though we never encountered these strings of letters as such. Once one poses the question, 'How is it possible to produce and understand sentences?', it appears as if we make plans (or generate 'new heuristic procedures') and put them into practice every time that we produce and understand ('new') sentences, even though these are familiar ones. That is to say, one is induced to suppose that there must be an 'unconscious' activity of processing information, of which we are not 'aware', for producing and understanding sentences, and to locate somewhere (in the brain!) this 'unconscious' invisible activity. Pylyshyn's (1973) argument, thus, seems to be based on the pseudo-problem of *how it is possible* to conceptualize 'novel' instances.

The idea of 'tacit' knowledge does not work in any case. Certainly, it is often difficult for us to formulate how to swim, how to ride a bicycle, how to use the word 'imagine', and so on, even though we know how to swim, how to ride a bicycle, how to use the word 'imagine', and so on. Do we know these things only 'tacitly'? Do we know them but only are *unaware* of them? If someone asks me whether I am aware of how to speak Japanese, my first language, I will not understand what the question means. Pylyshyn is right when he argues that participants in Kosslyn's experiments do not actually move an (imagined) speck on an image but rather imagine that a speck moves across a scene from one place to another. This, however, does not mean that they engage in processing information whose underlying representation is 'propositional'.[19]

Concluding Remarks

Of course, I do not deny that something must happen in the brain when one imagines things, and it is important to investigate neurophysiological *conditions* of imagination. What I have attempted to do is, however, to discuss imagination per se and to re-conceptualize it as an organizational property of the ongoing activity. We, as members of a culture,[20] know how to produce imagination in the course of organizing the activity we currently engage in. I have attempted to elucidate this knowledge we always already have.

Some may feel the following question is left open: how to explain the fact that 'more time is required to scan further distances across an image'?

There is nothing mysterious about this discovery. When we imagine a speck moving across a scene, we mostly imagine how it would move across a *real* scene (though, of course, you *do not have to* imagine a speck moving in this way). This is all. It is rather its psychological explanations that make it mysterious. Certainly, as opposed to what I suggested previously when I said that there is no discovery about one's own mental images, a participant in the experiment might find it surprising that she imagined as she *actually* did, that is, that she finely adjusted time to distances. According to Kosslyn, one might be surprised to learn something new about one's own mental images, and according to Pylyshyn, one might be surprised just because one's knowledge about the duration of a physical event is 'tacit'. However, the participant might be surprised at the result of the experiment, even though she is clearly aware of the general principle that further distances require more time. This surprise is, it seems, rather similar to the following one. I am very aware of the general principle that when one rides a bicycle one keeps one's balance by adjusting the handlebar and one's upper body to lurches of the bicycle, but if someone shows me with a slow-motion videotape recording of my riding how finely I adjusted the handlebar and my upper body to each lurch of the bicycle, I might be surprised at this.

Evidently, this surprise is neither because I learned a new thing nor because I became aware of what I had known only tacitly. Rather it is simply because I did not attempt to 'adjust the handlebar and my body in that way' in accordance with that 'general principle' when I rode a bicycle. What I attempted was just to 'ride a bicycle'. If I attempted to 'adjust the handlebar and my body in that way', I would have fallen over before deciding how to adjust them next and I would not have been able to ride a bicycle. Generally, even if one knows doing *P* (e.g. 'riding a bicycle') is *actually*, that is, from the observer's point of view, the same as doing *Q* (e.g. 'adjusting the handlebar etc.'), attempting to do *P* is not necessarily attempting to do *Q*. Therefore, if told after having done *P* that I *actually* did *Q*, I might be surprised. Similarly, it is because she did *not* attempt to 'adjust time to distances' in accordance with that 'physical theory' when imagining a speck moving that the participant in Kosslyn and his colleagues' experiment might be surprised if told after the experiment that she did actually finely adjust time to distance. What she attempted was simply to imagine a speck moving across a scene as she knows it would move across a real scene (even if she cannot formulate how she knows it moves across a real scene).[21]

Some may complain that they cannot find in this article a definition of what imagination or mental images *are*. Certainly, what I have attempted is to demonstrate in the analysis of a videotaped scene how imagination is organized in the actual course of activity. However, insofar as imagination is organized in a way appropriate to each current stage of the ongoing activity, imagination must have such diverse shapes, depending on all the particulars of the actual situation of the activity, that one cannot formulate a definition of its 'essence'. Imagination is not any event, process or state, or the like, that can be uniquely specified. If you try to take an 'essence' of imagination out of the sea of diversities by force, you will let this very imagination as we know it escape out of your grasp.

Notes

1. Psychological concepts such as imagination are basically polymorphous in their very character. I am not sure that the activity-related character of imagination covers the whole range of the concept. I am sure, however, that this is a very important characteristic that is very easily neglected in the psychology of imagination.
2. Of course, it is Pylyshyn who was most acutely critical of Kosslyn and his colleagues' claim that mental images are quasi-pictorial. He argues that *propositional* rather than depictive representations underlie the experience of having a mental image. For Kosslyn and his colleagues' attempt to defend their claim against Pylyshyn's criticism, see also Kosslyn (1981), Kosslyn, Pinker, Smith and Shwartz (1979), and so on. Kosslyn (1981) responds to Pylyshyn (1981), and Kosslyn et al. (1979) is countered by Pylyshyn (1979). See also Kosslyn (1995) for his summary of the debate between them. Pylyshyn is caught up in

the same confusion as Kosslyn insofar as he addresses the same question: what kinds of *mental entities* are involved in having a mental image? Both Kosslyn and Pylyshyn seem to cherish what Jeff Coulter (1979, 1989, etc.) calls 'the fallacy of reification', treating the expression 'mental image' as referring to some entities to be processed inside an individual. I will return to Pylyshyn's argument later.

3. For discussions on seeing per se in the same vein as this article, see Goodwin (1994, 1995, 1996), Goodwin and Goodwin (1996), Nishizaka (2000a, 2000b), and so on.
4. The relationship between imagining and remembering is undeniably important, and obviously the concepts 'imagining' and 'remembering' are related to each other. It might not be an accident that in many (probably most) experiments on mental images, participants are required to memorize figures of some kind. On the other hand, imagining is still conceptually quite different from remembering, and it is not even a species of the latter; imagining that a centaur runs on Santa Monica Boulevard is not any kind of remembering something, though it might be *conditioned* on the 'memory' of what centaurs are, what Santa Monica Boulevard is, and so on. Indeed, many students of mental images, including Kosslyn (1994, p. 324; Kosslyn et al., 1979, pp. 541–542; etc.), mention a variety of functional relations between mental images and short-term and long-term 'memory' (though I am suspicious of the usage of 'memory' in the literature).
5. For an overview of this debate, see Cornoldi and Logie (1996).
6. It is not so striking if Reisberg's (1996) following remark is taken into account: 'To serve as carriers of meaning, mental images must be more than "mere" pictures, and must, in particular, be unambiguous. It was this claim that we set out to test' (p. 165). The whole issue Reisberg and his colleagues address is to what extent mental images are more than 'mere' pictures.
7. The same point has been made by Gilbert Ryle (1949/1963) and Jean-Paul Sartre (1948).
8. Another recent focus of the psychology of mental imagery is on the cortical anatomy of mental imagery. For its overview, see Mellet, Petit, Mazoyer, Denis and Tzourio's paper 'Reopening the Mental Imagery Debate' (1998). They point out that 'there is a general consensus on the role of associative visual areas in mental imagery' and argue that this 'community of structure materializes the kinship between mental imagery and visual perception' (p. 132). I have nothing to say about the 'role of associative visual areas in mental imagery'. Which 'areas' facilitate and enable each other is a purely empirical issue to await empirical studies. However, what mental imagery per se is, what visual perception per se is, and what their relationship is are all conceptual issues, which precede every empirical issue in the sense that to specify which areas are related to visual perception, for instance, one must know in advance what visual perception *is*. Neurophysiology has contributed and will contribute much to the determination of functional relations between those cortical areas specified as facilitating mental imagery and visual perception, respectively, but this does not say anything about the 'kinship of mental imagery and visual perception'. Mental imagery and visual perception are *conceptually* related although they are not similar. This is true independent of any empirical findings. ('The language-

games employing these concepts [imaging and seeing] are radically different—but hang together' [Wittgenstein, 1967, § 625].)

9. We sometimes talk of 'an image *in* one's mind'. There is no problem with this expression as such, insofar as it makes sense in its actual use. Some might want to say that the expression is a metaphor. The point being made, however, is that, whether a metaphor or not, the expression has a radically different grammar from 'a picture *in* a frame'. The word 'in' in the expression 'an image in one's mind' does not provide the location of the image any more than the word 'in' of 'keys in one's possession' provides the location of the keys.
10. Commenting on the 'proof' of an external world by G.E. Moore, who held up each of his hands in succession, saying, 'Here is one hand and here is another; therefore at least two external objects exist; therefore an external world exists', Peter Winch (1958) remarks: 'Moore was not making an experiment; he was *reminding* his audience of something, reminding them of the way in which the expression "external object" is in fact used' (p. 10).
11. The phenomenological approach has been attempting to (re-)conceptualize imagination, too, by taking a rigorously descriptive attitude towards ways of experience. The general problem it raises is how it is that the world as we encounter it is as it is, or how it is that we experience the world as we do, and phenomenologists address the problem by reflecting on their consciousness and describing how we experience what we experience. As Hidé Ishiguro argues, though the phenomenological approach is radically different from the conceptual analytic one, which focuses on the use of language, both approaches advance similar claims. Indeed, Sartre's (1948) phenomenological exploration of imagination asks almost the same question as Ryle's (1949/1963) conceptual analysis, such as 'Does it make sense to say that . . . ?' (Ishiguro, 1966, p. 155; see also note 7) Another influential phenomenological study, by Edward S. Casey (1976), also asks the same kinds of questions, when, for example, he discusses differences between 'self-evidence' (of imagination) and 'incorrigibility', and so on. He argues, for example, that to be 'self-evident' is not to be subject to truth or error, such that to be self-evident is neither corrigible nor incorrigible. However, it seems that the phenomenological reflective description of experience may miss an important grammatical difference. It is true that what one imagines is 'self-evident' in Casey's sense, but this does not describe any *factual* conditions of experience, but rather provides a *normative* instruction for the use of the expressions 'I imagine', 'you imagine', and so on. (The import of this grammatical difference is the issue Ludwig Wittgenstein addressed in his remarks on colors, for example [Wittgenstein, 1990; see also Baker & Hacker, 1985].) Moreover, when Casey says 'the imaginative act-*cum*-presentation [is] experienced as genuinely self-evident' (p. 94), he seems to fall into what Sartre calls 'the illusion of immanence', that is, the misconception that the object of experience can be *in* consciousness. (Ishiguro argues that Sartre himself is often caught up in the same illusion in the course of his discussion.) In any case, what one imagines is 'self-evident' not because it is *experienced* as self-evident, but because saying 'I/you imagine something' logical-grammatically implies that it is self-evident to me/you.

    The approach I adopt here is most inspired by Harold Garfinkel's (1967) program. I attempt to explicate people's practices through which imagination is

normatively organized such that they imagine what to imagine in an appropriate way. Harvey Sacks (1992) investigated various interactional practices in the detailed analysis of tape-recorded 'naturally occurring' conversation. In the following analysis I owe much to his 'conversation analytic' investigations in the 1960s and the 1970s. See also Goodwin (1981) and Heath (1986) for the analysis of videotaped interaction in the direction Sacks established.

12. For the designers' own description of the system and the experiment they conducted with the system, see Suzuki and Kato (1995).
13. The above brief description of what the players are supposed to do in the fragments registers rationality of their practices insofar as it is intended as a description that participants themselves could provide about their own activity. It registers the sensibility, accountability, reportability, and so on, for participants of their activity. It is, on the other hand, a gloss for 'contingent ongoing accomplishments' of the organization of their activity as a 'rational' one. The following analysis is an attempt to investigate these 'organizational phenomena' as well. That is to say, to elucidate the concept of imagination and explicate practices for accomplishing imagination is to explore the way in which participants put to use imagination in ongoing accomplishments of rational properties of their activities. See Garfinkel (1967, ch. 1).
14. Symbols used in the transcripts and translations are:

    [Left-hand brackets, aligning vertically over two lines or (in Fragment 2b) extending over two lines, indicate the starting point of simultaneous talk.
    // A double slash indicates the point at which the next utterance starts.
    ::: Colons indicate that the prior sound is prolonged.
    (.) A dot in parentheses indicates an untimed micro-interval.
    ( ) Empty single parentheses indicate an inability to hear the utterance.
    . A period indicates a stopping fall in tone.
    ? A question mark indicates a rising intonation at the end of a phrase.
    ('P' in the word-by-word translations stands for 'particle'.)

    In the scene I focus on, the participants argue against each other so heatedly that it may be difficult to see what is going on only from a transcript of participants' utterances. Therefore, I present only a small fragment relevant to each analysis at once in the text, excerpted from the transcript of the whole scene.
15. I have analyzed another scene from the same—'AlgoBlock'—material elsewhere (Nishizaka, 2000a). There I pointed out that the participants in the game position themselves in what Adam Kendon (1990) called the 'F-formation', that is, a formation where a shared space is established and maintained from which parties immediately and readily reach for whatever objects their joint project may require they manipulate (p. 248). The point here, however, is not just the fact that there is a shared space for manipulations of objects, but rather the fact that participants see each other's orientations.
16. The expression 'They imagine possible routes of the submarine' is, in Garfinkel's (1967) terms, again a gloss for the ongoing accomplishment of rational properties of the current activity, insofar as it registers a sense the participants themselves could have about their own activity. See note 13.
17. I drew Figures 3 and 4 by tracing the original video frames.
18. In this context, it should be noted, however, that an actual route the submarine

takes after the 'play' button is pressed is also an imaginary one, insofar as the whole scene on the computer screen is imaginary. I owe to Wes Sharrock the comparison of imagined submarine routes with imagined air routes provided by airlines. On the other hand, there is a difference: whereas imagined air routes represent those real air routes, which are predetermined, each participant in 'AlgoBlock' draws imagined routes without any models given beforehand. Their drawing imagined routes is rather like a painter making a sketch using imagination. Saying 'using imagination' does not imply first having a certain image and then putting it onto a canvas. The painter, as it were, imagines with a paintbrush.

19. The well-known distinction between 'knowing-how' and 'knowing-that', made by Ryle (1949/1963), is relevant here. It should be noted, however, that neither 'knowing-how' nor 'knowing-that' is 'tacit' knowledge. The inability to formulate how to speak one's native language does not mean that one's knowledge of how to speak the language is 'tacit'. I, as a native Japanese speaker, know clearly very subtle differences between uses of various particles, even though I cannot formulate these differences. On the other hand, as I just argued, the (Chomskian) idea that we know how to speak the language but only 'tacitly' know the underlying grammatical theory (i.e. know that there are such and such general principles or rules) is incoherent and confused. Either (1) we do not know that theory (or those principles) at all, or (2) we are clearly aware of it, but, as I will argue shortly, the knowledge of it is not relevant to speaking the language.
20. I use here the term 'culture' very roughly. It does not mean anything bounded by a language, a life-style, a taste or the like. Rather, it only generally refers to competence for action and expression.
21. One may even be surprised to learn, through conceptual analysis, how 'seeing' and 'imaging' are conceptually related, even though one already knows how to use these expressions. This surprise, as Wittgenstein (1956) suggests about surprises in mathematics, reveals the fact that one did not command a 'perspicuous' view of the conceptual connection. See also Baker and Hacker (1985, p. 296). However, this surprise is different from the one I discuss in the text. To say that when one imagines a speck moving, one adjusts time to distances, or that when one rides a bicycle, one adjusts the handlebars, etc., registers an empirical and contingent, rather than conceptual, relation.

References

Baker, G.P., & Hacker, P.M.S. (1984). *Language, sense and nonsense*. Oxford: Blackwell.

Baker, G.P., & Hacker, P.M.S. (1985). *Wittgenstein: Rules, grammar and necessity*. Oxford: Blackwell.

Casey, E.S. (1976). *Imagining*. Bloomington: Indiana University Press.

Cornoldi, C., & Logie, R.H. (Eds.). (1996). *Stretching the imagination: Representation and transformation in mental imagery*. New York: Oxford University Press.

Coulter, J. (1979). *The social construction of mind.* London: Macmillan.

Coulter, J. (1989). *Mind in action*. Oxford: Polity.

Coulter, J. (1991). Is the 'new sentence problem' a genuine problem? *Theory & Psychology*, *1*, 317–336.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*, 606–633.

Goodwin, C. (1995). Seeing in depth: Space, technology and interaction on a scientific research vessel. *Social Studies of Science*, *25*, 237–274.

Goodwin, C. (1996). Transparent vision. In E. Ochs, E.A. Schegloff, & S.A. Thompson (Eds.), *Interaction and grammar* (pp. 370–404). Cambridge: Cambridge University Press.

Goodwin, C. (2000). Action and embodiment within human interaction. *Journal of Pragmatics*, *32*, 1489–1522.

Goodwin, C. (2002a). Time in action. *Current Anthropology*, *43*, 19–36.

Goodwin, C. (2002b). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where language, culture and cognition meet* (pp. 217–241). Hillsdale, NJ: Erlbaum.

Goodwin, C. (in press). The semiotic body in its environment. In J. Coupland & R. Gwyn (Eds.), *Discourse, the body and identity*. London: Palgrave/Macmillan.

Goodwin, C., & Goodwin, M.H. (1996). Seeing as situated activity: Formulating planes. In Y. Engestrom & D. Middleton (Eds.), *Cognition and communication at work* (pp. 61–95). Cambridge: Cambridge University Press.

Hacker, P.M.S. (1992). *Wittgenstein: Meaning and mind*. Oxford: Blackwell.

Heath, C. (1986). *Body movement and speech in medical interaction*. Cambridge: Cambridge University Press.

Ishiguro, H. (1966). Imagination. In B. Williams & A. Montefiore (Eds.), *British analytical philosophy* (pp. 153–178). London: Routledge.

Kendon, A. (1990) *Conducting interaction*. Cambridge: Cambridge University Press.

Kosslyn, S.M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review*, *88*, 47–66.

Kosslyn, S.M. (1994). *Image and brain*. Cambridge, MA: MIT Press.

Kosslyn, S.M. (1995). Mental imagery. In S.M. Kosslyn & D.N. Osherson (Eds.), *Visual cognition* (pp. 267–296). Cambridge, MA: MIT Press.

Kosslyn, S.M., Ball, T.M., & Reiser, B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 47–60.

Kosslyn, S.M., Pinker, S., Smith, G.E., & Shwartz, S.P. (1979). On the demystification of mental imagery. *The Behavioral and Brain Sciences*, *2*, 535–548.

Mellet, E., Petit, L., Mazoyer, B., Denis, M., & Tzourio, N. (1998). Reopening the Mental Imagery Debate: Lessons from functional anatomy. *Neuroimage*, *8*, 129–139.

Nishizaka, A. (2000a). Seeing what one sees: Perception, emotion and activity. *Mind, Culture and Activity*, *7*, 105–123.

Nishizaka, A. (2000b). The neglected situation of vision in experimental psychology. *Theory & Psychology*, *10*, 579–604.

Pylyshyn, Z. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*, 1–24.
Pylyshyn, Z. (1979). Imagery theory: Not mysterious—just wrong. *The Behavioral and Brain Sciences*, *2*, 561–563.
Pylyshyn, Z. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, *88*, 16–45.
Reisberg, D. (1996). The nonambiguity of the mental imagery. In C. Cornoldi & R.H. Logie (Eds.), *Stretching the imagination: Representation and transformation in mental imagery* (pp. 119–171). New York: Oxford University Press.
Ryle, G. (1963). *The concept of mind*. Harmondsworth: Peregrine. (Original work published 1949.)
Sacks, H. (1992). *Lectures on conversation* (2 vols.). Oxford: Blackwell.
Sartre, J.-P. (1948). *The psychology of imagination*. New York: Citadel.
Suzuki, H., & Kato, H. (1995). Interaction-level support for collaborative learning: AlgoBlock—an open programming language. *Proceedings of CSCL '95*, 349–355.
White, A.R. (1990). *The language of imagination*. Oxford: Blackwell.
Winch, P. (1958). *The idea of a social science and its relation to philosophy.* London: Routledge.
Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.
Wittgenstein, L. (1956). *Remarks on the foundation of mathematics*. Oxford: Blackwell.
Wittgenstein, L. (1967). *Zettel*. Oxford: Blackwell.
Wittgenstein, L. (1990). *Remarks on colour*. Berkeley: University of California Press.

ACKNOWLEDGEMENTS. Many thanks to Hiroki Aoki, Dom Berducci, Jeff Coulter, Nozomi Ikeya, Michie Kawashima, Shuya Kushida, Wes Sharrock and Brendan Wilson for their valuable and helpful comments on earlier versions of this article. For the analysis of the video fragments, I am deeply indebted to Chuck Goodwin's insightful and detailed comments. I am also grateful to Hideyuki Suzuki and Hiroshi Kato for their permission to use their rich and valuable material in this article.

AUG NISHIZAKA is Professor of Sociology at Meiji Gakuin University, Tokyo. His current research is concerned with reconsideration of psychological concepts from an ethnomethodological/conversation analytic point of view. ADDRESS: Department of Sociology, Meiji Gakuin University, Minato-Ku, Tokyo 108-8636, Japan. [email: augnish@soc.meijigakuin.ac.jp]