

The Neglected Situation of Vision in Experimental Psychology

Aug Nishizaka

MEIJI GAKUIN UNIVERSITY

ABSTRACT. The orthodox conception of 'visual perception' entertained in cognitive psychology is that retinal images inside eyeballs or impressions in the 'mind', which are very different from our actual visual experiences, are organized into a perception under an individual's skin. Even though this conception is caught up in profound 'grammatical' confusions about the concept of seeing or visual perception, nevertheless, it provides experimental psychologists with 'instructions' as to how to conduct experiments on 'visual perception'. In the main body of the paper an attempt is made to demonstrate how this orthodox conception is embodied in the actual course of interaction in psychological experiments. In the analysis of audio-visual recordings of an 'experimental' psychological experiment, I show that the actual course of interaction in the psychological experiment exhibits some distinct features, which are interactively organized so as to constitute the 'private' character of 'vision' in the experiment.

KEY WORDS: activity, interaction, psychological experiments, situation, visual perception

Introduction: Estranged Vision

Hans Wallach and D.N. O'Connell (1953) mention the classic problem for psychology of visual perception at the beginning of their well-known paper on 'The Kinetic Depth Effect':

The problem of how *three-dimensional form* is perceived in spite of the fact that pertinent stimulation consists only in *two-dimensional retinal images* has been only partly solved. Much is known about the impressive effectiveness of binocular disparity. However, the excellent perception of three-dimensional form in monocular vision has remained essentially unexplained. (p. 205, emphasis added)

In order to deal with this problem, they devised an experimental apparatus:

An object is placed between a punctiform light source and translucent screen and is rotated or turned back and forth. Its shadow is observed from

the other side of the screen. The shadow-casting object is placed as close to the screen as possible, whereas the distance between the light source and the object is made large. Owing to this arrangement isometric projection is closely approximated. The shadows of a great number of three-dimensional forms, solid or wire-edged, will be perceived as three-dimensional under these circumstances. The shadows of some forms will look three-dimensional *only* in such a moving presentation; that is, in none of the positions through which such a form passes during rotation will it cast a stationary shadow which looks three-dimensional. (p. 206)

The authors call this effect the 'kinetic depth effect.'

Wallach and O'Connell presuppose that, when one sees something, first two-dimensional images are provided on one's retina, and then in certain circumstances those images effect a three-dimensional perception.¹ Insofar as their basic problem is 'how three-dimensional form is perceived in spite of the fact that pertinent stimulation consists only in two-dimensional retinal images' and their experiments are designed to solve it, to them, the screen on which two-dimensional shadows are cast is an artificial retina.

However, the presupposition that the perception of three-dimensional form is constructed from two-dimensional retinal images has a well-known difficulty, which was noted even by Descartes: whereas shadows on a screen, the artificial retina, are seen by people facing the screen and the kinetic depth effect is perceived by them, who is it who sees the images on the 'natural' retina and who is it who gains the depth effect? This difficulty, that is, the need for a 'seer' of retinal images inside a person, is what Anthony Kenny (1984) called 'the homunculus fallacy'.

These 'private' retinal images, which are seen by one's mind's eye, are the deprived images, that is, deprived of the third dimension. Starting from those private retinal images, the question for those researchers should be: how can these images be organized such that a three-dimensional perception is effected? It should be noted, however, that when seeing something, one is never first given private/deprived retinal images. As Merleau-Ponty (1964) says, one cannot call depth the 'third dimension':

First of all, if it [depth] is a dimension, it would rather be the first one; there would be no certain forms nor planes unless I could define at what distance from me their different parts are found. However, such a first dimension that contains all the other dimensions is not one dimension. . . . Depth as conceived this way is rather the experience of the reversibility of dimensions, the experience of a global 'locality' where everything is there, from which height, largeness and distance are only abstracted, or the experience of the voluminosity which means precisely that there is a thing. (p. 65)

Our lived space is not one that extends along three axes (x , y , z) at right angles to each other, that is, the space that we encounter in mathematics textbooks. Our lived space is not constructed in the way that one dimension

is added to a two-dimensional plane. Indeed, we encounter two-dimensional planes in our lived space. For example, images on paper are two-dimensional. But it should be noted that when one sees a planar image, one sees at the same time that it lies on the table at a certain distance from one's face. In other words, one would not be able to see a plane if one did not see at the same time the depth in which it were embedded; just as one can see white letters on a blackboard only if one also sees the blackboard at the same time. Planar images on paper are to be seen only in the voluminosity into which we have always already (*immer schon*) been thrown along with things.

To conceptualize vision as a result of a third dimension being added to a two-dimensional plane is to estrange vision *to* the geometrical space. The private/deprived retinal images, which are alienated *from* the third dimension, are the *consequence* of alienated vision, rather than its source.

This said, however, the aim of the present paper is not to further pursue this philosophical issue of the 'homunculus fallacy', which has been tackled by many philosophers and psychologists from diverse points of view (see, e.g., Bickhard & Richie, 1983; Hacker, 1987; Hyman, 1989; Katz, 1983). It is rather to demonstrate the social organization of the activity of performing a psychological experiment, conducted by psychologists with the erroneous conception of 'visual perception'. Though the basic conception of 'orthodox' cognitive psychology is caught up in fatal conceptual confusions, the fact that a distinct activity is conducted with the conception is as such a phenomenon to be investigated in its own right. In this respect, my interest here is rather sociological than philosophical or psychological.

Incidentally, the gist of the 'homunculus fallacy' argument is that absurdities result from extended applications to *parts* of an organism, for example its sensory organs, nervous system, 'mind', or whatever, of mental predicates that should only be applied to a *whole* organism in their ordinary usage (cf. Kenny, 1984, p. 130). This implies that human action, perception, or whatever, cannot be analyzed into any double construct of 'neat' behavior and meaning to be attached to the behavior, or 'neat' data (retinal images, sense impressions, or whatever) and interpretation to be provided to the data. In these terms, the basic conception of orthodox psychology of visual perception can be formulated more generally as the conception that given (retinal, neural, mental, or whatever) images or sensations are processed, interpreted, organized, or whatever, into perception under an individual's skin.²

The Distinctness of Doing a Psychological Experiment

Performing a psychological experiment is a distinct activity. My demonstration focuses upon the distinctness of the activity. By 'distinctness', however,

I do *not* mean the *uniqueness* of the activity, that is, a set of necessary and/or sufficient conditions for the activity, but a set of any nameable features of the activity that are discernibly related to the activity. The focus is especially upon those features of the social organization of the activity that are discernibly related to the basic conception in psychology.

In demonstrating how the activity is organized, I will attempt to analyze some fragments from videotaped 'experiments', but it should be borne in mind that the analysis is not aimed at empirically verifying any hypotheses. It is rather an attempt to explicate our knowledge, or *sociologic*, of the organization of the activity. Indeed, I, as a sociologist, have no experience with conducting a psychological experiment. However, I think, I can be a competent 'subject'³ of an experiment if I have a chance (and actually I have had one). Every competent member of our society must be potentially a competent subject of an experiment. The demonstration here attempts to explicate this competence, which every competent member is *normatively* expected to have. (Note that the 'competence' here is not an empirical but a normative concept.)

In what follows, no attempt will be made to criticize Wallach and O'Connell's *hypothesis* on the kinetic depth effect. Wallach and O'Connell attempted to specify the conditions under which a two-dimensional plane looks three-dimensional.⁴ I do not have anything to say against their claim of the specified conditions under which this took place. Nothing is mysterious in their experiments; *you* can see two-dimensional images on the screen, and these images look to *you* three-dimensional under the conditions *you* can specify. Confusions break out only when they conceive of the screen before viewers as representing the latter's retina (or any 'mental' screen whatever) and their whole series of experiments as elucidating the nature of visual perception of three-dimensional form, which is supposed to be achieved through internal processes under each individual's skin. It is this move that invites some conceptual confusions.

It should be noted here that the claim that seeing is not achieved physiologically or psychologically through processing images or sensations is not an *empirical* one. The claim cannot be verified nor falsified by any (empirical) observations. Indeed, it must be true that something is 'cast' on retinas, for example, but the expression 'two-dimensional retinal images' is still in confusion. As mentioned above, seeing a two-dimensional plane implies the living and seeing in the voluminosity where the viewer belongs together with, and at some distance from, the plane. However, it is impossible for retinal images to be seen in this way, because it is impossible to imagine something (someone?) behind the retina that sees those images in that voluminosity (which would imply that it has a body, from which the images are at some distance). In this respect, 'retinal images' cannot be two-dimensional. The question of whether 'two-dimensional retinal images' are first provided is not an empirical matter, but rather only a conceptual

nonsense. (Certainly, if you detach a retina from a body and project something on it through a 'natural' or artificial lens, you will obviously obtain a two-dimensional image on it, which is two-dimensional precisely *to you* who see it.) It is rather a *conceptual* matter, concerning logical grammars of the concepts 'plane', 'two-dimensional', 'vision', 'see', and the like.⁵

In the main body of this paper, I will analyze some fragments from visually recorded sessions of an 'experimental' experiment, which imitates Gunnar Jansson and Gunnar Johansson's (1973) experiment.⁶ Their original experiment proceeds like this:

The proximal stimuli in the form of outline quadrangles changed in small perceptually-subliminal steps between two extreme forms, the first one being always a square. After the other extreme form had been reached the form change was reversed until the original square form was regained. . . . Six stimuli were studied. (Jansson & Johansson, 1973, p. 322)

Those stimuli, photographed frame by frame, were projected on the screen before the 'subject' (see Figure 1).

The subject was told that the aim of the experiment was to study how some changes of quadrangles were perceived. . . . An instruction card with the three main motion categories described in a few words was handed over to the subject, and the meaning of them was demonstrated by cardboard models. (p. 324)

Those categories are 'rotation,' 'bending' and 'stretching'; the 'subject' was told to choose from them the one that the stimulus presented to them seemed to fall under. 'It was emphasized that other types of percepts could also be expected and that, in that case, a verbal description of the percept was wanted' (p. 324).

It should be noted that Jansson and Johansson do not, as Wallach and O'Connell did, attempt to specify the conditions on which 'three-dimensional' forms can be 'perceived' from 'two-dimensional' shadows;

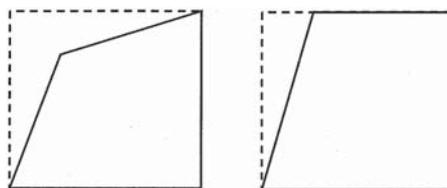


FIGURE 1. Two of the six stimuli in Jansson and Johansson's experiment (Jansson & Johansson, 1973, p. 323, reproduced with permission by G. Jansson and Pion Limited, London). The stimuli used in our 'experiment' are black figures against a white background.

rather they attempt to demonstrate a perceptual preference ‘for rotary motion over bending motion, and of bending motion over two-dimensional form change’ (p. 321). To Jansson and Johansson, ‘two-dimensional’ images are not necessarily given prior to a ‘three-dimensional’ perception; both are effects of some moving shadows. However, to them, too, some ‘projected’ images are independent variables, and these images *look* differently in accordance with their different motions. That is to say, those images presented before ‘subjects’ are also here intended to represent their private/deprived images (retinal, neural, ‘mental’, or whatever), which are supposed to be processed by them under their skin.

The following fragment is excerpted from one of the videotaped sessions. In the ‘experiment’, subjects are told to classify the images on the monitor into the three categories provided on the instruction card: rotation, bending and stretching. Then, subjects are supposed to, so to speak, ‘process’ the information given through the ‘stimuli’ by themselves so as to organize their ‘perception’ of changing forms. In the following fragment, a subject is trying to classify an image shown to her:

- #1 [PP4; original in Japanese; simplified]⁷
- 1 S: *Nanka* [Kind of] it came down this way and went back
 2 → °*mitaina* [or something like that] (°), °*uuuun*° *de* [then],
 3 wh- what happened, *dore daro* [which one is it] [hh hhh.
 4 E: [*Maa* [Well],
 5 you can say it’s none of them *desu kedomo*

After describing the movements of the image, at line 2, she, the subject, starts to think about how to classify it by herself. First, the expression ‘*uuuun* [mm]’ marks her thinking and ‘*dore daro* [which one is it]’ marks that she has not yet found the right classification. Second, she first looks at the computer monitor when she describes the movements of the image, but then she looks down at the instruction card in one of her hands and puts the other hand on her mouth:

#1a⁸

Monitor	Instruction Card
S: went back	° <i>mitaina</i> (°), ° <i>uuuun</i> ° <i>de</i> , wh- what happened, <i>dore daro</i>
	└──────────────────────────────────┘ ((S puts her hand on her mouth))

Thus she is ‘thinking’ by herself to process the information given by the stimulus on the computer screen. On the other hand, not only is she thinking by herself, but she is *observably* thinking by herself; her ‘thinking’ is transparent to others and, indeed, can be used as a resource for collabor-

actively organizing the current interaction. In fact, at line 5 the ‘experimenter’ responds to her ‘thinking’ and suggests that she could just say none of the three categories can apply if she finds it difficult to categorize the image according to the instruction card.

Of course, there is nothing mysterious in the fragment. Simply an activity is conducted in interaction. However, right here is suggested the (apparently paradoxical) distinctness of the activity of performing a psychological experiment: observably *individual* ‘processing’ (e.g. classifying) of given information is lodged in the *social* organization of a joint activity. In what follows, I will elaborate on this point by showing how the basic conception of psychology is embodied and instantiated by this distinctness.

The Situated Character of Vision

Before going into the detailed analysis of the above fragment, some words are in order about a basic feature of vision, which is very important and, nevertheless, tends to be neglected when discussing vision or visual perception.

Of course, when one sees something, light strikes one’s retinas and the retinas are stimulated (something is ‘cast’ on them), and then the stimuli are transmitted to the brain. This is an undeniable fact. However, it is important to bear in mind that these physiological processes are *not* seeing or vision per se. They are just conditions enabling one to see, and do not as such constitute seeing. It should be noted that, as Ryle (1949/1963, pp. 143–146) suggests, seeing is not a process of any kind but an achievement. It very often happens that we mistake something for something else in seeing, like this: though I thought I saw a *real* oasis, it has turned out to be just a mirage. That is to say, seeing in the ordinary sense is an achievement that *might not* have been accomplished.

As Kenny (1984) remarks,

... the normal way to discover whether an organism sees is not just to study its eyes, but to investigate whether its behaviour is affected by changes of light and colour, etc. Consequently, an explanation of seeing must be an explanation not only of the acquisition and storage of information, but also of what makes the containing of the information into knowledge—i.e. its relation to behaviour. (p. 130)

It is not when we specify particular physiological processes (or any processes or events inside an individual whatever) that we can say someone sees something, but rather under some *circumstances*, including his or her behavior, which are publicly accessible; for example, we can say someone sees a mirage when he or she says ‘There is a mirage!’, pointing in some direction where others also see a mirage, without starting to run in that direction with an empty canteen in his or her hand, or even when he or she,

shouting 'Oasis!', starts to run in the direction where others see a mirage of an oasis. When we say someone sees a mirage, we ascribe to him or her not some specific physiological processes, but some abilities and/or motives; for example, we ascribe the ability to refrain from starting to run in that direction though there can be seen something like an oasis over there, or a motivation that drives him or her to run in that direction though there is nothing over there. Thus, insofar as the actual situation of seeing is neglected, even if the details of the neural system are greatly clarified in the future, it will not help specify further what seeing is (beyond just what conditions make seeing possible). *Seeing is essentially situated.*

The concept 'situation' here refers to an indefinite set of those circumstances (actual or potential) available to more than one person and relevant to their current activities. The situated character of vision, therefore, implies that vision is a *public* achievement.

The claim 'I see/saw *X* can be criticized by others on reasonable grounds. First, it is criticizable by examining whether *X* is/was out there. Second, it is criticizable with reference to its propriety. If you *can* see someone being strangled in an apartment opposite to yours and, nevertheless, you only see two distinct moving objects being in touch with each other, or only the cord on the person's neck, your seeing may be regarded as morally inappropriate; that is, your competence for social life may be doubted, though what you claim you see may be a 'fact' or a 'real' object in a way. Seeing is an achievement that is open to not only cognitive, but also moral criticisms from others.

On the other hand, at least in some cases, even 'appearance' can be also a public achievement, though the claim 'It appears to be *X*' or 'It looks (like) *X*' does not imply that it is really *X*, as the claim of seeing does.⁹ At least in some uses of 'appear' or 'look', appearance is also open to others' criticisms. First, for example, someone may look like a swindler to me. However, when I say this to my wife, who also looks at the same man, she may disagree with me, saying, 'He looks rather honest'; she may reject my claim by presenting some grounds such as his manner of dress, his tender look, and so on. Then I may argue with her on grounds that support my original claim. Second, if everybody that I come across looks to me like a swindler, I may be morally criticized ('You are too suspicious,' 'You have a warped mind,' or the like) and, here again, my moral competence for social life may be doubted.

What is important here is that seeing and appearance are not only cognitive, but also normative phenomena. Some may say that the so-called 'surface perception', that is, perception of depth, is rather mechanical: we are forced to see one thing on another, rather than the other way around, for example a cord on a man's neck rather than a man's neck on a cord (or a band partly occluded by a man's neck?). This may be true. As suggested, however, at least insofar as human cognition is concerned, 'surface percep-

tion' is only achieved in the context of seeing various things and/or events within ongoing activity. That is why it is morally inappropriate to see only a cord on a man's neck when we should see rather a man being murdered.

Furthermore, it should be noted that we sometimes see an empty box. When we see this kind of thing, we do not see a box with a feature of 'emptiness', but see that there is nothing in it. It is in a sense surprising that we see there is nothing, because there are an indefinite number of 'nothings' in the world and, nevertheless, only in some cases is it reasonable to say we see there is nothing. For example, we usually do not say we see there is nothing between the top of my computer monitor and the ceiling of the room, though it is true that nothing is there. Generally speaking, what we see is embedded in our activity, or our 'form of life'. When voting last time, I saw a quadrangular space surrounded by black lines on a slip of paper, inside which I wrote down a candidate's name; but I did not see those lines were slightly crooked at several places, even if they were actually so. Thus, we should see what we are normatively expected to see within our current activity or our 'form of life'. That voluminosity we live and see in is essentially normatively tinged.

In psychological experiments, too, what subjects see is embedded in the context of their ongoing activity, that is, doing an experiment. However, at the same time, their conduct is organized to be 'individual', as we saw above. The 'thinking' of the subject in Fragment 1 is observably 'individual'. Furthermore, the 'vision' that is attributed to subjects can be said to be 'private', as will be shown shortly, in the sense that their reported 'vision' is deprived of a chance to publicly be focused on in the subsequent development of interaction, where it might be able to be further criticized, evaluated, and so on. Note that I use the term 'private' here to generally refer to some distinct observable features of subjects' conduct in psychological experiments, rather than features in principle inaccessible to others. The aim of the present paper is to show how the orthodox conception of psychology is embodied in those observable features. In what follows, in the analysis of some videotaped fragments of (experimental or practice) psychological experiments, I will show how the 'individual' or 'private' character of subjects' conduct or 'vision' is socially or interactively organized in accordance with the basic conception of orthodox psychology.¹⁰

A Conception Embodied in Interaction

A Conception as 'Instructions'

As mentioned above, the basic conception of orthodox psychology is not an empirical hypothesis. It cannot be falsified on the basis of any empirical examinations. It is not because the conception is correct, but just because the

confusions it is caught up in are not empirical but conceptual. On the other hand, this conceptually confused basic conception, nevertheless, can provide the framework in which empirical psychological research should be conducted, that is, it can constitute 'instructions' as to how to perform psychological experiments on 'visual perception'. That is to say, orthodox psychologists put their erroneous conception of vision to use as a framework or 'instructions' by which they organize their experiments on 'visual perception', yet they have actually gained some *empirically* testable findings on 'vision'; that is, the conception can work quite well for their purposes. This conception is instantiated by and embodied in the actual course of interaction within psychological experiments on 'vision'. This is, as stated above, a phenomenon that should be investigated in its own right. The question to be asked here is then: how is this orthodox conception embodied in and through the actual development of the activity of performing a psychological experiment?¹¹

Indeed, this conception is embodied in the design of experiments. For example, subjects' vision is deliberately detached from the situation in which they are co-present with others, through various kinds of experimental devices.¹² Here, however, the focus will be rather upon the *in situ* design of the development of interaction in a psychological experiment.

Cutting-Off of Follow-Ups

In each session of the videotaped 'experimental' experiment mentioned above, the experimenter reads out what is contained in his prepared manual. What he, the experimenter, reads is an arranged translation of Jansson and Johansson's (1973, p. 324) description of what they told their 'subjects' (which was cited above). However, though no description of what he tells his 'subjects' after the general instruction is found in his manual, the interaction between the two participants thereafter has a prominent basic pattern. Here are two examples:

#2 [PP2; original in Japanese; simplified]

1 E: Describe your impressions how they look,¹³ with words, please.

2 S: Yes.

3 E: (That's the way.) *Dewa*, [Then,] the first one.

4 ((Two cycles of a change of a quadrangle))

5 E: This is it.

6 S: Just the upper-left caved in

7 E: Yes.

8 S: The original quadrangle was compressed this way or something.

9 E:→*Ah*, [Oh,] deformation?

10 S: Right.

11 E: It was being deformed *to* [you say]

12 S: Like being deformed

13 E:→I see, I see. okay. *Jaa, tsugi* [Then, next] ()

#3 [PP4; original in Japanese; simplified]

- 1 E: Well *jaa* [then] let's go to the first one.
- 2 ((An angle of a quadrangle moves a little bit towards the center and
2a moves back to the original position))
- 3 E: What does this look like?
- 4 S: This, like bending
- 5 E:→*Ah*, [Oh,] looks like bending.
- 6 S: Yes.
- 7 E:→Yes. *Jaa*, [then,] *tsugi wa* [next]

In both cases (subjects are different), the experimenter (E) marks the starting point of the main body of each experiment with '*dewa*' or '*jaa*', or 'then' (at 3 in #2 and at 1 in #3). After letting a stimulus on the computer screen move, he encourages the subjects to choose a category from the instruction card now in the subjects' hands, only by marking out the end of the movement of a stimulus (at 5 in #2) or explicitly in a question format (at 3 in #3). Then the subjects (S) give an 'answer', that is, a description of the stimulus or a category from the card (at 6–8 in #2 and at 4 in #3).

There are some prominent features in these fragments: First, when the experimenter repeats or rephrases the subjects' answers (at 9 in #2 and at 5 in #3), he marks their answers with a token that marks out that new or valuable information having been provided, or what Heritage (1984) calls a 'change-of-state token' ('*Ah*'; cf. Schiffrin, 1987). Second, generally speaking, it is usually expected that when one gains markedly new information, it will be shown in one way or another what is new about it. It might sound strange if someone were to ask another a question, for example, 'What are you doing now?', and then, in responding to the answer he receives, were only to say, 'Oh, you are' and leave without any other comments or follow-ups. However, in fact, the experimenter makes no follow-ups, whereby the value to him of the information in the subjects' answers could be made transparent. Third, not only are there no follow-ups, but any possible further development from each answer is expressly *cut off* by the experimenter's utterance '*Jaa, tsugi* [Then, next]' (at 13 in #2 and at 7 in #3). Note that the expression '*Jaa, tsugi* [Then, next]' marks out the next development of interaction as 'next', that is, severed from the preceding one. That is to say, at the exact place where he sees there might be a follow-up, he does not allow it to go on but rather proposes that they move on to a 'next' stimulus.

Perhaps he should not have used any 'valuable information' markers at all or even have said anything after subjects had given their answers. However, there is another case where experimenters say expressly something like 'Then, next'.

The following fragment is from a videotaped session of a practice experiment, conducted by two undergraduate psychology students after Experiment 1 from Kosslyn, Ball and Reiser (1978). In accordance with the

manual that the experimenters (E1 and E2) prepared in advance, they instruct the subject (a student of another department: S) to memorize ‘accurately’ an array of three letters spaced on paper, image it in her mind after memorizing it, and then ‘focus’ on the letter ‘A’ when experimenters say, ‘Please focus on “A”’, for instance. The subject is then told to ‘move a black speck’ in her head from the ‘focused’ letter to the target letter as fast as possible, by the shortest route and without letting the speck leap. When she ‘arrives at’ the target, she has to press the ‘yes’ button, or she has to press the ‘no’ button if the ‘target’ given by experimenters is not included in the array.

#4 [Image Scanning Experiment: original in Japanese]

- 1 E1: Please imagine the positions of A, B and C.
- 2 ((Pause))
- 3 E1: OK? Please focus on C.
- 4 ((Pause))
- 5 S: °What should I say°
- 6 E2: O? ‘Yes.’
- 7 E1: Only ‘yes’ is fine.
- 8 S: Yes.
- 9 E1: A.
- 10 S: ((Presses a button.)) ()
- 11 E2: That’s OK.
- 12 S: That’s it?
- 13 E1: Yes, it is.
- 14 S: O. ((Nods.))
- 15 E1:→*Dewa tsugi ni iki masu*. [Then, we will move on to the next.]
- 16 Please focus on B

Those experimenters basically read their prepared manual. The manual reads like this: ‘Please focus on C. (Yes) “A” “Please focus on B” (Yes) “F” “Please focus on A” (Yes) “B” Please open your eyes’ (original in Japanese; quotation marks are exactly reproduced). Note that no mention is made about subjects’ responses on those buttons in the prepared manual. That is, these responses have no place in the interaction format set up (in accord with that orthodox conception) in the manual. It should also be noticed that the phrase ‘*Dewa tsugi ni iki masu* [Then, we will move on to the next]’ is not included in the manual. Now attention should be drawn to the sequential context in which this phrase is used.

The trial reproduced as Fragment 4 is the very first one of the exercise before the main experiment, and, as is easily seen, the subject has trouble in proceeding in accordance with those instructions that were given by the experimenters. After, at line 5, she has forgotten what she was told to say when she has focused on the named letter, she says, at 10, something (inaudible), where she is supposed to press a button silently. E2 confirms the

appropriateness of S's conduct at line 11. Note that the subject's request for reconfirmation at line 12 ('*Korede ii no?* [That's it?]) sounds like an indication that she feels a lack of something in only silently pressing a button. The phrase at 15 ('Then, next') is related to this feeling of a lack. It indicates explicitly that there is nothing more to be dwelt on within the 'current' trial and that they should move ahead right now. The phrase is thus designed *in situ* to implement the format in the manual, that is, to deprive the subject's response on a button of any chance to (further) be focused on in interaction. Indeed, in the main experiment after the exercise, they do not use this phrase ('Then, next') and just go on to saying 'Please focus on . . . ' when the subject presses a button.

The 'Then, next' construction is thus a device for cutting off what is possibly interactionally relevant, that is, a possible subsequent development of interaction, in which, for example, participants might have a chance to elaborate on the value of the information extracted from subjects' responses. Though the experimenter in Fragments 2 and 3 claims that he has gained valuable information from the subjects' descriptions of how the images on the computer screen look, the information is deprived of that interactional relevance that is brought about by his claim, and therefore deprived of a chance to be focused on publicly in the subsequent development of interaction, where it can further receive criticisms, justification, evaluation, and whatever.¹⁴ In this way the 'private' character of the subjects' 'vision' is achieved within the actual development of interaction.

This is an example, of course. A particular device for achieving the 'private' character of 'vision' in psychological experiments is not always used. Moreover, usually no such device is used except under special conditions. The point to be made here is that in one way or another, depending on the interactional contingencies, the orthodox conception of 'visual perception' is embodied in the actual development of interaction in a psychological experiment.

The Interactive Constitution of 'Private' Information Processing

The subject in Fragment 1 was involved in processing of the information that is given on the computer screen. Her information processing can be said to be 'private' *in a sense*. What information was extracted, which category will be selected and how the processing is conducted are observably left to the subject's decision. The following fragment is an extended reproduction of Fragment 1:

#5 [PP4; original in Japanese; simplified]

1 S: *Nanka* [Kind of] it came down this way and went back

2 →^o*mitaina* [or something like that] (°), ^o*uuun*^o *de* [then],

- 3 wh- what happened, *dore daro* [which one is it] hh^v [^ hhh.
 4 E: [Well,
 5 you can say it's not none of them^v [^ *desu kedomo*
 6 S: [*Demo, haa*
 7 ((Pause))
 8 E: °Uh°
 9 S: *Uuuun*, still, only this side came this way, and, and went
 10 →back=*demo* [but], (.) °*ah, demo, ah* (.) ben-° bending is not it
 11 *desu kedo*
 12 E: *Ah*
 13 S: °something like that *desu*.°
 14 E: *Ah*, I see. Next is . . .

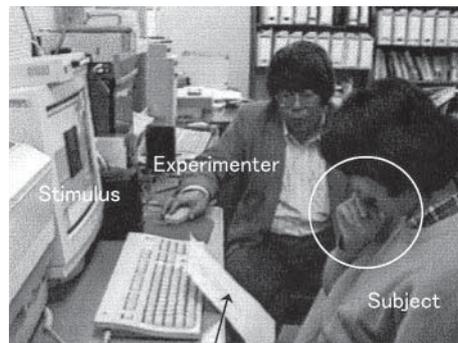
There are two very similar segments here (ll. 1–3 and ll. 9–11), where the subject first tries to describe the movements of the image she has just seen and then does ‘think’ about which category applies here. Certainly, it may not be clear from the above transcript that at line 10 she is doing thinking, but it will become obvious if her bodily movements are considered together. At line 10 she makes very similar bodily movements to those at line 3. In both cases, while she gazes straight at the computer monitor when she describes the movements of the image, she gazes down at the instruction card in one of her hands and she brings her other hand to her mouth or nose while doing ‘thinking’.

#5a [ll. 1–3; identical to #1a]

Monitor Instruction Card

S: went back ° mitaina(°), ° uuuun° de, wh- what happened, *dore daro*

((S puts her hand on her mouth))



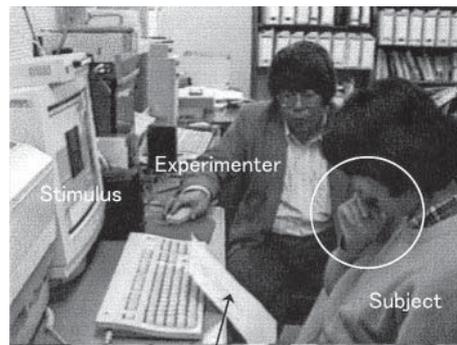
Instruction Card

#5b [ll. 9–11]

Monitor Instruction Card

S: and, and went back=demo [but], (.)^o ah, demo, ah (.) ben-^o bending└──────────────────┘
|

((S puts her hand on her nose with her eyes closed))



Instruction Card

Note, moreover, that in Fragment 5, all over the subject's utterances the only part she designs to be explicitly addressed to the experimenter is 'bending is not it *desu kedo*' at line 11. First, this is the only part in which the subject uses a polite expression ('*desu*'); that is, the most part of her utterances are not designed for someone whom she has just met for the first time. Second, at line 10, just when she mentions 'bending', she starts to bring up her head away from the card towards the computer monitor and raises her upper body; it looks as if she is getting out of the 'thinking' position. Indeed, the experimenter responded to her with a 'valuable information' token after she utters the polite expression '*desu kedo*'.¹⁵

Of course, the subject's conduct of 'thinking by herself' does not constitute a *uniquely* psychological activity. Nevertheless, the construction of the conduct can still be said to embody the orthodox conception of 'vision'. On the one hand, information extracted from the image on the computer screen, the category being applied to it and the processing of the information are observably left to the subject's decision. On the other, this 'private' character of the 'process' is interactively constituted precisely as such in a very subtle way. In what follows I will show how the interaction in Fragment 5 is organized such that the 'private' character of the subject's information processing is jointly accomplished.

Now I will focus on the subject's utterance at ll. 1–3 first. It can be seen that after the part 'and went back' the utterance is coming to an end. Not only does the phrase '*mitaina* [or something like that]' mark out conventionally that a 'transition-relevance place' of a turn at talk (a place where another

person than the current speaker may take a turn at talk [Sacks, Schegloff, & Jefferson, 1974]) is coming or even has come, but also, after uttering ‘and went back’, the subject acutely lowers the volume of her voice (as indicated by degree signs), which also marks out that the transition-relevance place is coming. (Moreover, while uttering ‘*mitaina*’, she starts to lower her upper body as if retreating from the speakership.) Certainly, there are no gaps in her utterance at ll. 1–3, but it should be noted that if a transition-relevance place is markedly coming, and if the recipient, that is, the possible next speaker, does not start a next utterance at once, the ‘delay’ in response is perceivable on the side of the possible next speaker (for this point, see Jefferson, 1973). For example:

#6 [SFD; original in Japanese; simplified]
 Yasu: Such and such village in such and such county *mitaina*,
 $\overset{\vee}{\text{sooiu}}$ $\overset{\wedge}{\text{yoona}}$ [like this] (°)
 Akio: $\left[\begin{array}{l} \text{Hoooooooooooo} \\ \text{oo} \end{array} \right]$
 Heita: $\left[\begin{array}{l} \text{Jaa (.) anoo. . .} \end{array} \right]$

In this fragment, when some tokens (using the verbal expressions ‘*mitaina*’ and ‘*sooiu yoona*’ and lowering the volume of his voice) mark out that a transition-relevance place is coming,¹⁶ Akio starts to respond and, perhaps in responding to this response, Heita starts his utterance, even before Yasu stops his utterance.

Compared with this fragment, at line 2 in Fragment 5 the absence of a (normatively expected) response on the side of the experimenter is now clear. Right here, the experimenter makes a prominent bodily movement:

#5c [ll. 1–3]

((E leans over the instruction card in the subject’s hand))

S: and went back $\overset{\vee}{\text{mitaina}}$ (°), $\overset{\vee}{\text{uuuun}}$ de

((S gazes away from the monitor at the card,
and lowers her upper body))

wh- what happened, which one is it

((S puts her hand on her mouth))

The experimenter starts to lean over the instruction card in the subject’s hand at almost the same time as the subject starts thinking by uttering the sound ‘*uuuun*’ (which is conventionally used when doing ‘thinking’). He

does two things with this movement. First, he encourages the subject to continue to categorize the image with a category on the instruction card, by enhancing his orientation to the card. Second, by doing so, he provides an account for the absence of his response to her utterance: that he should refrain from speaking up because the subject is supposed to classify the image by herself. He passes the opportunity to speak presented to him *and* provides an account for the absence of his taking a turn at talk at once with one bodily movement. It is precisely at this moment that the subject begins doing thinking with '*de*' (which, like 'and' used at the beginning of an utterance, marks both connection to and disconnection from the preceding utterance).

Thus, the 'private' character of the subject's information processing is an interactional product; it is collaboratively accomplished by both participants in the interaction. First, as was suggested above, though the subject's thinking by herself ('What happened, which one is it?') is not designed to be *addressed* to the experimenter, it is still designed to be *heard* by him. She does not just think by herself, but *does* thinking by herself in front of the experimenter. Thinking by herself is used here as an interactional resource for letting their interaction go on without any explicit disturbance.¹⁷ It both is occasioned by *and* contributes to the actual development of interaction. Second, the subject is, so to speak, urged by the experimenter's conduct not to stop until she has finished classifying the given image by herself.¹⁸

Now turning to the second segment of the subject's talk (ll. 9–11): here again, after trying to describe the movements of the image on the computer screen, the subject does thinking until she finally proposes an answer mentioning a category on the instruction card ('bending'). In this segment, the experimenter's conduct initially appears to be of no interactional relevance unlike in the first one. However, a close examination of the segment reveals that this is not true. When she starts doing thinking by uttering '*demo* [but]' at line 10, this word ('*demo*') *perceivably* 'latches' onto, that is, follows with no interval, the preceding one, as designated by an equals sign (=); that is, the subject's proceeding to doing thinking is achieved in a marked way. If the subject provided the experimenter with an opportunity to take a turn at talk after some attempts to describe the image on the computer screen, as she actually did before, the experimenter *would* refuse to take a turn to talk again, as he actually did before. That 'latching' is a way to cope with this possibility. The subject retreats to thinking by herself in order to avoid in advance the perceivable absence of an expected response that would possibly result otherwise. The latching *makes* any response from the recipient unexpected or irrelevant and thereby prevents any possible absence of a response from being perceived. Thus, here again, the 'private' character of the subject's conduct can still be said to be interactively achieved.

In sum, the whole process of the subject's classifying the given image (or information processing) is organized so as to assume the 'private' character. Indeed, this very character is interactively, and therefore publicly, organized through a kind of negotiation between the co-participants. In this way, here again, the participants in the experiment jointly accomplish the embodiment of the basic conception of orthodox psychology in the actual course of interaction.

Conclusion

We have seen how the 'vision' being ascribed to subjects in a psychological experiment was organized to be 'private'. However, as stated above, nothing is mysterious in the psychological experiment. It is real human beings who see given images and process the information obtained from the computer screen, not 'homunculi' inside the brain or mind. It is images on the screen that they see, not those on a natural retina or anything inside an individual. The participants actually visually perceive 'two-dimensional' images on the screen, the instruction card, drawings on it, co-participants' faces and bodily movements, subjects' doing thinking, and so on. Their seeing these things is entirely embedded in their ongoing activity; their seeing must be relevant to the development of the activity. For example, when and how the experimenter should ask 'What does the image look like?' is contingent on his seeing an image on the screen *and* his seeing the subject looking at the image. A subject's providing an answer to the question depends on his or her seeing the image on the screen then. Seeing, or visual perception, is thus an achievement embedded in the ongoing activity the participants are jointly engaged in, rather than any process or event lodged within the nervous system or mind; seeing provides visual information to be processed within a social activity, rather than consists in a 'mental' activity of 'information processing' within the nervous system. Seeing is a structural feature of an 'activity system' in that it is organized through and organizes the ongoing social activity.

Thus, very 'ironically', the way the participants see things within the activity of carrying out an experiment in accordance with the basic conception of orthodox psychology sharply contradicts this very conception.

As said, however, the 'private' character of the 'vision' that is attributed to subjects is not deniable as a *phenomenon*. Certainly, the 'vision' does not have the private nature as the orthodox conception conceives it. It is not any process, activity or event under an individual's skin. But the 'vision' is still organized to be 'private' in a sense (though what subjects report is not what they *perceive* but how what they perceive on the screen *looks*; note,

however, that the orthodox conception [confusedly] conceives the whole process the subject is involved in—from seeing images on the screen through mentioning a category of the instruction card—to represent a ‘process of vision’). What has been shown is that this ‘private’ character is interactively, publicly accomplished. There is no paradox here. Just as a two-dimensional plane is lodged in the voluminosity where we live with others, so the ‘private vision’ being attributed to subjects in psychological experiments is lodged in the interactional, public and normative context of a distinct activity.

Appendix: Original Japanese Transcripts and Word-by-Word Translations

Additional symbols used in the original transcripts are:

::: Colons indicate that the prior sound is prolonged.
 hhh .hhh A series of h's marks an outbreath (hhh) and an inbreath (.hhh).

Abbreviations used in the word-by-word translations are:

IR = interrogative
 JD = judgmental
 NG = negative
 P = particle
 PL = polite

#1 (the first part of #5; see below)

#2

E: ...Dooyuufuuni mieru ka no (.) inshoo o, ano kotoba de
 how look IR as to impression P uh words with

setsumee shite kudasai.
 explain do please

S: Hai
 yes

E: (°toiukoto desu°.)
 that's it JD+PL

((Two cycles of a change of a quadrangle))

S: Kore wa hidari ue ga hekkonde iru yoo(h)ni
 this P left upper P dented be like

shika(h) (.) hidari ue ga
 only left upper P

- E: Ee.
yes
- S: Motomoto no shikakui katachi kara (.) kooiufuuna kanji de
the beginning at quadangular form from like this about P
asshukusarete iku 'tseiuka
be compressed be going to or
- E: (.) Aa, henkee?
oh deformation
- S: Soo desu ne.
so JD+PL P
- E: Henkee shite [iru to
deformed be being P
- S: [Henkee shiteiru yooni
deformed be being like
- E: Haa haa, haa haa. °Wakari mashita°. .hhhh Jaa, tsugi ()
yeah see did+PL then next
- #3
- E: Kore wa:: nani [ni:: mie:: masu ka::?]
this P what P look like PL IR
- S: [Kore wa:: : : (.)] wankyoku::,
this P bending
- E: A, wankyoku ni mie masu::.
oh bending P look like PL
- S: Ha[ji
yes
- E: [Hai °Jaa° (.) tsugi wa::
yes then next P
- #4
- E1: Me wa renshuu ga owaru made tojita mama de ite
eye P practice P end till close keep
kudasai.
please
A, (.) B, (.) C no ichi o atama no naka de imeeji shite
of location P head P inside P image do
kudasai.
please
- E1: Iidesu ka? C ni shooten o awasete kudasai.
ok IR on focus P adjust please
- S: Nante yuu no?
what say P

E2: E? 'Hai'.
what? yes

E1: ['Hai' dake de ii de [su.
yes only P ok JD

S: [Hai. [Hai.
yes yes

E1: A

S: ()

E2: Sore de (.) ookee.
that P ok

S: Kore de ii no?
this P ok P

E1: Hai, sore de:.
yes that P

S: A,
o

E1: Dewa, tsugi ni iki masu. B (.) ni shooten o-
then next to go PL on focus P

#5

S: Na- nanka nanameni, koo:: taorete kite iru (.) 'tteyuuka,
kind of diagonally this way down come have or

mekurete kite, modotta (.) °mitaina° (), °uu::n°, de, do- doo
rolled up got went back like that and what

natta n'daroo, dore daroo. hh [hh
happened I wonder which I wonder

E: [Maa, dore demo nai,
well anyone P NG

S: Uu::n

E: ['tte itte mo ii 'n [desu kedomo::],
P say even if ok JD+PL P

S: [demo:: (.) haa]
but yeah

E: ((pause)) °Ee::°
yes

S: Uu::n, ((pause)) yappari, kocchi dake ga koo::,
still this only P this way

(.) hidari gawa no hashi- koko dake ga
left side P end this part only P

koo:: kite iru, (.) mata (.) mata modotta =demo:(.) °ah, demo
 this way ·come have again again went back but uh but

ah (.) wan-° wankyoku to iukanji demo nai n'desu kedo
 bending P like either NG JD+PL P

E: Ah
 oh

S: °Sonna kanji° (.) desu.
 like that seem JD+PL

E: °Ah::° hai ((a cough)) wakari mashi ta .hhhh Kondo wa ...
 oh yes see PL did this time P

Notes

1. This presupposition is still entertained in the mainstream cognitive sciences. For example, see David Marr's influential book (Marr, 1982). Gibson (1979/1986) cites plenty of experimental grounds for suspicion against this orthodox conception. While Gibson's 'unorthodox' theory of visual perception provides an illuminating insight into visual perception, he seems still not to pay sufficient attention to the situated character of seeing, with which this paper is concerned. That is to say, he conceives vision as a process or activity that takes place somewhere, that is, extracting information in the signals reaching receptors. For criticisms of this (rather orthodox) supposition of Gibson's, see Coulter (1995), Coulter and Parsons (1991) and Sharrock and Coulter (1998).
2. Perhaps, for the purpose of the present paper, the basic conception of orthodox psychology should be formulated in such a way as to include all the conceptions that do not pay enough attention to the 'situated', 'public', 'social activity embedded' (etc.) character of (human) vision. See below.
3. The *Publication Manual of the American Psychological Association* (Fourth Edition) suggests that the word 'subject' be replaced by 'participant' because of the former's pejorative implications (American Psychological Association, 1994, p. 49). In this paper, in accord with the tradition of sociological studies of human interaction, I use the word 'participant' as a general term for parties to an interaction and refer to *both* the experimenter and the 'subject' as 'participants'. Indeed, in view of psychological experiments as a form of human interaction, experimenters are also 'participants' in interaction. Moreover, as it will turn out, I focus on the interactional accomplishment of some specific aspects of 'being a subject' (not 'being just another participant') in a psychological experiment that seem to be related to the origin of the term 'subject'. For these reasons, I would rather use the word 'subject' in this paper.
4. Their hypothesis is formulated like this, for example:
 . . . a shadow line must undergo both a displacement and a lengthening or shortening in order to produce a kinetic depth effect. Both these changes must be given together. A change in length alone is not sufficient to produce a reliable kinetic depth effect. (Wallach & O'Connell, 1953, p. 210)

It should be noted that they specify here on what conditions a two-dimensional plane *looks* to the subjects to be three-dimensional.

5. Button, Coulter, Lee and Sharrock (1995) show that one of the fundamental mistakes in cognitive sciences consists in pursuing *empirical* studies to solve *grammatical* or conceptual problems.
6. There are a couple of reasons why I have chosen Jansson and Johansson's experiment instead of Wallach and O'Connell's one, whose 'experimental' sessions have been videotaped, too. First, Jansson and Johansson describe in their paper the procedure of their experiment in some detail. In Wallach and O'Connell's paper, no clues are found about what exchange of talk occurred between the experimenter and the subject. Second, all that was available to us in our experimental experiments were images on a computer monitor; behind the monitor there were not objects which cast their shadows, as in Wallach and O'Connell's case, where you can say safely, '*They see* a rotating object,' instead of only saying, '*It looks to them* to be rotating,' though actually Wallach and O'Connell discuss when the shadows *look* three-dimensional to the subjects.
7. All fragments in the main text are a simplified translation of original Japanese transcripts. Some particles whose original forms and/or sequential positions are relevant to the analyses are reproduced in italics. The original transcripts of Fragments 1–5 appear in the Appendix at the end of this paper. Symbols used in the fragments are:

- [] A left-hand bracket and a right-hand bracket, stretching over two lines, indicate the starting point and the ending point of simultaneous talk, respectively.
- (1.6) A number in parentheses indicates in seconds and tenths of a second the length of a time interval within an utterance or between utterances.
- (.) A dot in parentheses indicates an untimed micro-interval.
- () Empty single parentheses indicate an inability to hear the utterance.
- ◦ Items between two degree signs have a noticeably lower volume than the surrounding talk.
- = An equals sign indicates 'latching', that is, no interval between the end of a prior and start of a next part of talk.
- . A period indicates a stopping fall in tone.
- ? A question mark indicates a rising intonation at the end of a phrase.

Some more words about my data are in order here. I contacted some cognitive psychologists to obtain some data on psychological experiments. They explained to me that they were anxious to avoid any uncontrolled influences upon their experiments (e.g. they kept their laboratory dark during the experiments to restrict subjects' field of vision to the computer screen on which the stimuli appeared). Indeed I was not permitted to videotape their 'real' experiments. That is to say, the more relevant to my research experimenters are, that is, the more 'orthodox' they are, the more difficult it is to obtain visual recordings of their 'real' experiments.

8. 'Monitor' on the solid line indicates that she gazes at the computer monitor during the time period designated by the length of the line; 'Instruction Card' indicates that she gazes at the instruction card in her hand during the designated time period.

9. I neglect here all the subtle differences between appearance and look that J.L. Austin (1962) describes. See also Coulter (1979) and Hacker (1987).
10. Certainly, the term ('private') is used ambiguously in this paper. But this ambiguity is rather intended. On the one hand, when, for example, used in the phrase 'private retinal images', the term refers to being in principle inaccessible to others in a strong ('philosophical') sense, that is, something logically unintelligible. On the other hand, here it refers to distinct phenomena. The point to be made is, as mentioned repeatedly, that, while the idea that vision, information processing, and so on, are 'private' in the former (non-)sense is simply in confusion, some *phenomena* that can be said to be 'private' in the latter (more ordinary) sense are organized in accordance with that idea.
11. Michael Lynch (1991) says:

Gibson argues that the psychology of vision based on such experiments [in which the subject is precluded from using what Gibson calls 'ambient' and 'ambulatory' vision] is erroneous, but for our purposes this 'error' constitutes a kind of 'effective historical consciousness' [Gadamer]. The orthodox theory is embodied in technology and in the literary figure of a human actor categorically bound and disciplined in accordance with the technological complex. (p. 57)

Perhaps, we should say 'the orthodox conception' instead of 'the orthodox theory'. In any case, it can be said further that the orthodox *conception* of visual perception is embodied in the actual course of interaction where those technologies are employed by those actors. See also Lynch and his colleagues' studies of 'instructed' action (e.g. Lynch, in press; Lynch & Jordan, 1995; Lynch & Law, 1999), where they discuss how formal 'instructions' are implemented in various kinds of activities, such as telephone survey, molecular biology, bird-watching, and so on, from the perspective opened up by Garfinkel (1967).

12. For example, in experiments of visual perception, the subject often has one of his or her eyes fixed such that artificial 'private retinal images' are produced. Jansson and Johansson (1973) in their experiment use the arrangement like this:

The subject sat on the opposite side of the screen looking monocularly through a large collimator lens at focal distance from the screen. . . . The peripheral parts of the collimating lens were occluded by a black screen close to the lens. The opening around the central part of the lens was rectangular (40 cm × 20 cm) and the subject's eye was at a distance of 35–40 cm from the lens. From this position the angle subtended at the eye by the visual field was about 60 deg × 30 deg, and by the stimuli in square form 2.5 deg × 2.5 deg. . . . (p. 323)

13. It should be noted, again, that the experimenter asks the subject what it *looks* like, in spite of Jansson and Johansson's original description of the procedure; they say 'the subject was told that the aim of the experiment was to study how some changes of quadrangles were *perceived*' (p. 324, emphasis added).
14. See Garfinkel (1996), in which he respecifies Gestalt phenomena from the ethnomethodological point of view and cites Doug Maynard's observation: 'In a classroom context, working with gestalt figures, perception and its production cannot be separated from public descriptions that students and professors produce and attend to as joint courses of action' (p. 17).

15. At line 3, the subject invites the experimenter to interrupt her by using a laughing token at the end of the utterance. However, she does not address the utterance to him, but only invites him to come in; note that she constructs the utterance in a rough form ('*daro*', as contrasted to '*deshoo*').
16. Note that the expression '*mitaina*' is as such ambiguous in that we can sometimes stop utterance with it in a natural way, while it is 'originally' used with a noun phrase following it, like this: '*neko mitaina inu* [a dog that looks like a cat]'. In these terms, it may be said that the place where Akio started, rather than the one right after '*mitaina*', was the earliest possible one where a next speaker could start.
17. See Goffman's (1981) argument on 'self-talk', which is often designed to be heard by others co-present with the speaker, though it is not addressed to others. For the 'thinking face' gesture as an interactional resource for organizing a distinct activity, see M.H. Goodwin and C. Goodwin (1986).
18. 'Thinking by oneself' is not an entirely 'mental' activity. The Goodwins' study of searching for a word (cited above) provides a detailed demonstration on this point. See also C. Goodwin (1994) and (1996) for the interactive accomplishment of classification within distinct activities.

References

- American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.
- Austin, J.L. (1962). *Sense and sensibilia*. Oxford: Oxford University Press.
- Bickhard, M.H., & Richie, D.M. (1983). *On the nature of representation: A case study of James Gibson's theory of perception*. New York: Praeger.
- Button, G., Coulter, J., Lee, J., & Sharrock, W.W. (1995). *Computers, mind and conduct*. Cambridge: Polity.
- Coulter, J. (1979). *The social construction of mind*. London: Macmillan.
- Coulter, J. (1995). The informed neuron: Issues in the use of information theory in the behavioral sciences. *Minds and Machines*, 5, 583–596.
- Coulter, J., & Parsons, E.D. (1991). The praxiology of perception: Visual orientations and practical action. *Inquiry*, 33, 251–272.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Garfinkel, H. (1996). Ethnomethodology's program. *Social Psychology Quarterly*, 59, 5–21.
- Gibson, J.J. (1986). *The ecological approach to visual perception*. New York: Erlbaum. (Original work published 1979.)
- Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96, 606–633.
- Goodwin, C. (1996). Practices of color classification. *Ninchi Kagaku*, 3(2), 62–82.
- Goodwin, M.H., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62, 29–49.
- Hacker, P.M.S. (1987). *Appearance and reality*. Oxford: Blackwell.
- Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J.M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 299–345). Cambridge: Cambridge University Press.

- Hyman, J. (1989). *The imitation of nature*. Oxford: Blackwell.
- Jansson, G., & Johansson, G. (1973). Visual perception of bending motion. *Perception*, 2, 321–326.
- Jefferson, G. (1973). A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in Closing Sequences. *Semiotica*, 9, 47–96.
- Katz, S. (1983). R.L. Gregory and others: The wrong picture of the picture theory of perception. *Perception*, 12, 269–279.
- Kenny, A. (1984). *The legacy of Wittgenstein*. Oxford: Blackwell.
- Kosslyn, S.M., Ball, T.M., & Reiser, B.J. (1978). Visual image preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 47–60.
- Lynch, M. (1991). Laboratory space and the technological complex: An investigation of topical contextures. *Science in Context*, 4, 51–78.
- Lynch, M. (in press). The living text: Written instructions and situated actions in telephone surveys. In D.W. Maynard, H. Houtkoop, N.C. Schaeffer, & H. vander Zouwen (Eds.), *Standardization and tacit knowledge: Interaction and practice in the survey interview*. New York: Wiley.
- Lynch, M., & Jordan, K. (1995). Instructed actions in, of and as molecular biology. *Human Studies*, 18, 227–244.
- Lynch, M., & Law, J. (1999). Pictures, texts and objects: The literary language game of birdwatching. In M. Biagioli (Ed.), *The science studies reader* (pp. 317–341). London: Routledge.
- Marr, D. (1982). *Vision*. New York: Freeman.
- Merleau-Ponty, M. (1964). *L'oeil et l'esprit*. Paris: Gallimard.
- Ryle, G. (1963). *The concept of mind*. Harmondsworth: Peregrine. (Original work published 1949.)
- Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
- Sharrock, W., & Coulter, J. (1998). On what we can see. *Theory & Psychology*, 8, 147–164.
- Wallach, H., & O'Connell, D.N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, 45, 205–217.

ACKNOWLEDGEMENTS. Many thanks to Chuck Goodwin, Shuya Kushida, Mike Lynch, Hiroki Maeda, Satoyuki Morita and Naoki Ueno for their valuable comments on an earlier version of this paper. I am also very grateful to Dom Berducci for his helpful comments.

AUG NISHIZAKA is Professor of Sociology at Meiji Gakuin University, Tokyo. His work has been focused on a reconstruction of classical topics in social and human sciences from the ethnomethodological/conversation analytic point of view. Among his recently published articles are 'Doing interpreting within interaction: The interactive accomplishment of a "henna gaijin" or "strange foreigner"' (*Human Studies*, 22, 1999) and 'Seeing what one sees: Vision, emotion and activity' (*Mind, Culture and Activity*, 7, 2000). ADDRESS: Meiji Gakuin University, Minato-ku, Tokyo 108–8636, Japan. [email: augnish@soc.meijigakuin.ac.jp]