

# Reading the Implicit Mind from the Body

*Makio Kashino, Makoto Yoneya, Hsin-I Liao, and Shigeto Furukawa*

### Abstract

Recent studies in cognitive science have repeatedly demonstrated that human behavior, decision making, and emotion depend heavily on the *implicit mind*, that is, automatic, involuntary mental processes that even the person herself/himself is not aware of. We have been developing diverse methods of decoding the implicit mind from involuntary body movements and physiological responses such as pupil dilation, eye movements, heart rate variability, and hormone secretion. If mind-reading technology can be made viable with cameras and wearable sensors, it would offer a wide range of usage possibilities in information and communication technology.

*Keywords: man-machine interface, physiological signals, eye movement*

### 1. Introduction

In daily life, people often infer other people's feelings and intentions to some extent—even if they are not expressed explicitly by language or gesture—by taking the appearance of the person and the situation into account. This ability, often referred to as *mind reading*, is an essential characteristic that supports smooth communication among people. A majority of current information and communication technology (ICT) devices, on the other hand, do not function without receiving explicit commands, which are entered using predetermined methods such as typing, pressing buttons, using one's voice, and making specific gestures. If ICT devices had a mind-reading ability, the relationship between such devices and users would be more flexible and natural. Ultimately, users would not be aware of the existence of ICT devices. In other words, mind-reading technology would make ICT devices transparent to users.

### 2. Reading the mind from the body

Mind-reading technology has been a topic of extensive research in recent years. In fact, remarkable progress has been made in brain-computer interfaces

(BCIs), which decode a person's brain activity to identify the content of the person's consciousness, such as a category of a perceived object, or a button to be pressed among multiple alternatives. Our approach, however, is essentially different from BCI in two aspects.

The first difference concerns the method of measurement. In BCI, brain activity is measured using such technologies as electroencephalograph or functional magnetic resonance imaging, whereas in our mind reading approach, we measure physiological changes on the body surface, including eye movements, pupil diameter changes, heart rate variations, and involuntary body movements. These signals can be measured with relatively simple devices such as a camera or surface electrodes. In contrast, BCI requires large-scale specialized measurement equipment. At present, measurement of body surface signals is not completely unconstrained, meaning that the movement of the person being measured is somewhat restricted. However, it will be even less constrained and more transparent in the near future, when sophisticated wearable sensors are developed.

The second difference concerns the decoding target. BCI tries to categorize the content of consciousness, such as types of perceived visual objects (e.g., a

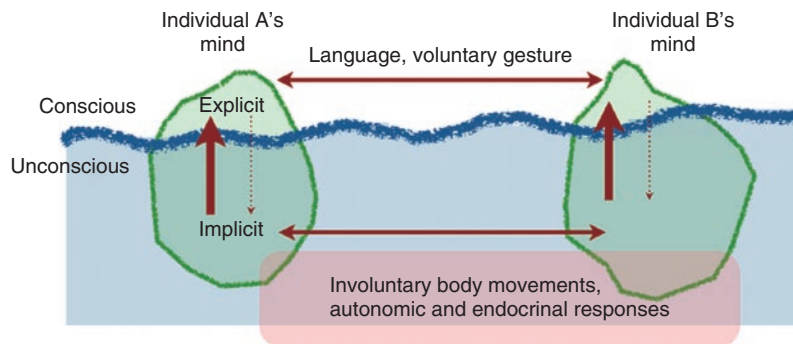


Fig. 1. Illustration depicting implicit and explicit minds.

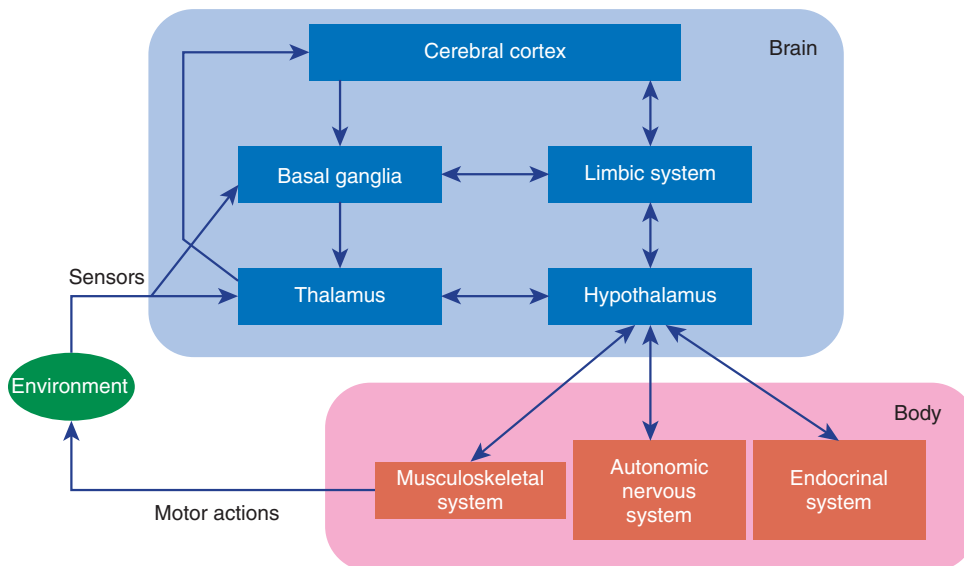


Fig. 2. Loops of the brain, body, and environment.

face vs. a house) and a button to be pressed (e.g., left or right). The system learns the statistical correspondence between the categories and brain activity patterns in advance. Based on the learning, it judges which category the observed pattern belongs to. However, it should be noted that consciousness is only a fraction of the whole *mind*. Recent findings in the field of cognitive science repeatedly demonstrated that human behavior, decision making, and emotion depend not only on conscious deliberation, but they also depend heavily on the implicit mind, that is, automatic, fast, involuntary mental processes that even the person herself/himself is not aware of (**Fig. 1**) [1]. This implicit mind is the target of our mind reading.

Many experimental studies have shown strong interactions between the implicit mind and the body. In a sense, they are inseparable, like two sides of a coin. Such a tight relationship reflects the complex loops of the brain, body, and environment (**Fig. 2**). If some event happens in the environment, the states of one's body such as the autonomic, endocrinal, and musculoskeletal systems change so that the person can react to the event appropriately. The information about the event is also sent to the cerebral cortex, where the event is recognized through complicated information processing. These changes in body states start prior to, and thus deeply affect, the processing in the cerebral cortex. This is why people often *lose control* of their mind and body, no matter how well

they understand what to do. For example, if they had to make a speech in front of some eminent people, they might inadvertently make some uncharacteristic mistakes due to extreme nervousness.

Moreover, in interpersonal communication, unconscious body movements of partners interact with one another, creating a kind of resonance. The resonance, in addition to explicit language and gesture, may provide the basis for understanding and sharing emotions (Fig. 1). An experiment we conducted recently demonstrated that the unconscious synchronization of footsteps between two people who had met for the first time and were walking side by side for several minutes enhanced the positive impressions that they had of each other. The mind is, in one aspect, a dynamic phenomenon that emerges through the interaction mediated by bodies. Thus, measuring the body surface instead of the brain not only has practical merit, but is also essential in mind reading. The next section introduces an example of our experiments.

### 3. Reading the familiarity and preference for music from the eyes

#### 3.1 Measuring microsaccades

As the saying goes, “The eyes are more eloquent than the mouth.” In the context of mind-reading technology, gaze direction has been used extensively as an index of visual attention or interest. However, what is reflected in the eyes is not limited to mental states directed to or evoked by visual objects. We are studying how to decode mental states such as saliency, familiarity, and preference for sounds based on the information obtained from eyes, namely, a kind of eye movement called a microsaccade as well as changes in pupil diameter. In the experiment, the measurement is conducted using a high-precision eye camera (sampling rate = 1000 Hz, spatial resolution  $< 0.01^\circ$ ), installed in a downward frontal position to the participant, whose head movement is restricted by a chin rest (Fig. 3). Although technical problems remain to be solved before we can achieve completely unconstrained measurement in the real world, it has been shown that the eyes can provide more diverse information than previously thought.

Microsaccades are small, rapid, involuntary eye movements that typically occur once every second or two during a visual fixation task (Fig. 4(a)). We have revealed a previously unexplored relationship between auditory salience and features of microsaccades by introducing a novel model of eye-position control [2]. In short, the presentation of a salient (i.e.,



Fig. 3. Experimental setup to decode familiarity and preference for music from the eyes. Eye movements and pupillary responses of participant while she listens to a tune are measured using an eye camera.

unusual or prominent, and therefore easily noticeable) sound among a series of less salient sounds induced a temporal decrease in the damping factor of microsaccades, which is an indicator of the accuracy in position control, and a temporal increase in the natural frequency of microsaccades, which is an indicator of the speed of position control (Fig. 4(b)).

#### 3.2 Studying pupillary responses

One more thing we focus on is pupillary responses. The primary function of the pupils is to control the amount of light entering the retina, just as the diaphragm of a camera does. However, pupil diameter is also modulated by emotional arousal and cognitive functions such as attention, memory, preference, and decision making (Fig. 5(a)). This is because pupil diameter is controlled by the balance of sympathetic and parasympathetic nervous systems, and reflects, to some extent, the level of neurotransmitters that control cognitive processing in the brain. We have demonstrated that pupil dilation, as well as microsaccades, can be used as a physiological marker for certain aspects of auditory salience [3]. Temporary pupil dilation occurs when a salient sound is presented among a series of less salient sounds (Fig. 5(b)). Such pupil dilation responses depend on various factors including acoustic properties, context, and presentation probability, but not critically on voluntary attention to sounds.

We applied these basic findings in some attempts to estimate a listener's familiarity and preference for a tune based on the features of microsaccades and pupillary responses while listening to the tune. In

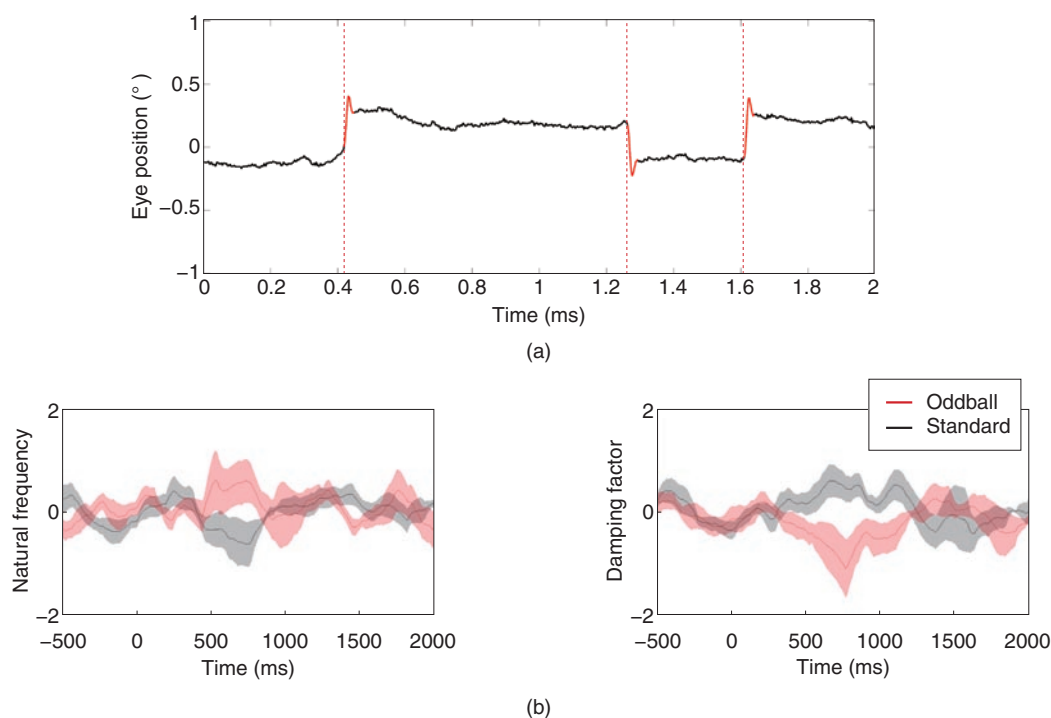


Fig. 4. (a) Example of microsaccade measurement. Microsaccades are indicated by red lines. (b) Time evolution of microsaccade parameters (natural frequency and damping factor) in response to oddball (rare and thus salient) and standard (common and not salient) sounds. The horizontal axis represents time from onset of sound. Each sound was 50 ms in duration.

addition to the physiological signals obtained from the eyes, we also analyzed the acoustic/musical properties of the tune. For example, we have developed a novel *surprise* index, which represents the extent of unpredictability of the musical data at a given moment in a tune, given the data up to that point. We consider this index useful because a typical tune consists of regularity (predictability) and deviation from it (unpredictability or surprise), and the balance between the two seems to be one of the critical factors that contribute to familiarity and attractiveness of the tune.

At NTT Communication Science Laboratories Open House 2014, we conducted a demonstration in which each participant listened to one of 15 tunes from various music genres including classical, rock, and jazz for 90 seconds. The participant's familiarity and preference ratings for the tune were then estimated based on 12 features of microsaccades and pupillary responses, together with several features of the tune including surprise. Prior to the demonstration, the decoding system had learned the mapping between those features and subjective ratings (7-point

scale each) of familiarity and preference for 23 participants. In the demonstration, the differences between the actual and estimated ratings were 2 or smaller in more than 80% of the trials for nearly 200 participants. (Note that this was only an informal demonstration and not a rigorous test.) The demonstration was designed to estimate subjective (that is, not implicit) ratings, but could be extended to implicit mind (or behavior), once objective behavioral data are available, such as which tune each participant decided to buy and when and how many times the tune was played back.

#### 4. Future directions

In addition to the work described in this article, we are conducting diverse lines of research concerning the responses of the brain and autonomic nervous systems, hormone secretion, and body movements. In one such project, we developed a method to measure the concentration of oxytocin (a hormone considered to promote trust and attachment to others) in human saliva with the highest accuracy to date (in

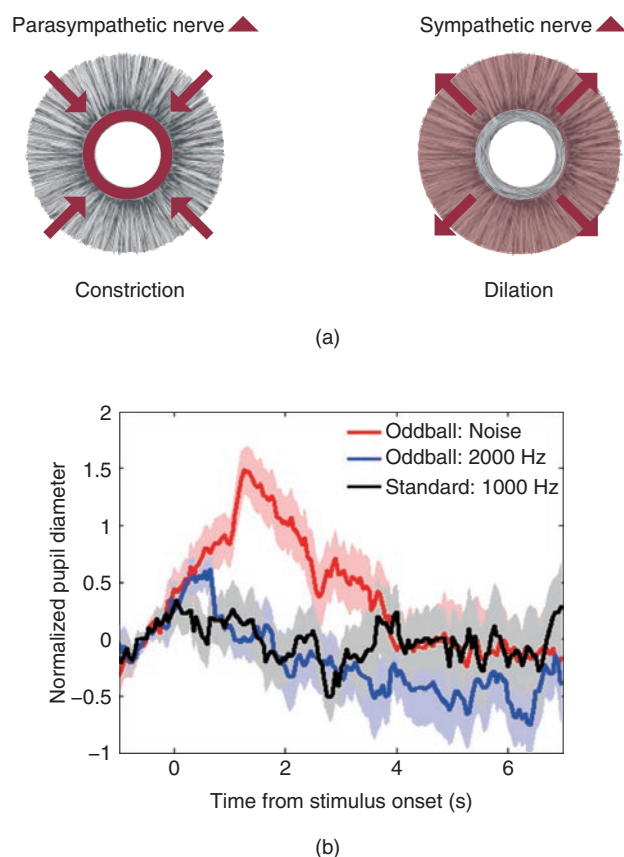


Fig. 5. (a) Relationship between autonomic nervous system activity and change in pupil diameter. (b) Pupil diameter changes in response to oddball and standard sounds. Duration of each sound was 50 ms.

collaboration with Prof. Suguru Kawato of the University of Tokyo). This has enabled us to identify a physiological mechanism underlying relaxation induced by music listening. Listening to music with a slow tempo promotes the secretion of oxytocin, which activates the parasympathetic nervous system, resulting in relaxation [4]. Analysis of hormone concentration is not fast enough for direct use in ICT devices. However, when the relationship between physiological signals that are quickly measurable and the relevant hormone concentration is revealed by

laboratory experiments, the knowledge will be beneficial for mind reading. In evaluating the quality of video or audio, for example, it would be possible to capture the differences in quality of experience that are not apparent in subjective ratings.

As wearable sensor technology advances, mind reading will be applied more widely. It is especially attractive in sport-related areas. For example, measurement of heart rate, respiration rate, and muscle potential at various parts of the body using sensors woven into underwear would make it possible to report mental and physical states of players during a game, or to develop effective training methods that combine monitoring and sensory feedback.

The study of reading the implicit mind has just begun. For the moment, basic research is necessary to understand the mechanisms of the complex loops of mind and body. Needless to say, careful consideration should be given to ethical issues such as privacy and safety.

### Acknowledgments

Part of this research is supported by CREST from Japan Science and Technology Agency and SCOPE (121803022) from Ministry of Internal Affairs and Communications.

### References

- [1] D. Kahneman, "Thinking, Fast and Slow," Farrar, Straus and Giroux, 2011.
- [2] M. Yoneya, H.-I. Liao, S. Kidani, S. Furukawa, and M. Kashino, "Sounds in Sequence Modulate Dynamic Characteristics of Microsaccades," Proc. of the 37th Annual MidWinter Meeting of the Association for Research in Otolaryngology, PS-606, San Diego, CA, USA, February 2014.
- [3] H.-I. Liao, M. Yoneya, S. Kidani, M. Kashino, and S. Furukawa, "Human Pupil Dilation Responses to Auditory Stimulations: Effects of Stimulus Property, Context, Probability, and Voluntary Attention," Proc. of the 37th Annual MidWinter Meeting of the Association for Research in Otolaryngology, PS-599, San Diego, CA, USA, February 2014.
- [4] Y. Ooishi, H. Mukai, K. Watanabe, S. Kawato, and M. Kashino, "The Effect of the Tempo of Music on the Secretion of Steroid and Peptide Hormones into Human Saliva," Proc. of the 35th Annual Meeting of the Japan Neuroscience Society, P2-h01, Nagoya, Japan, September 2012.





### Makio Kashino

Senior Distinguished Scientist/Executive Manager of Human Information Science Laboratory, NTT Communication Science Laboratories.

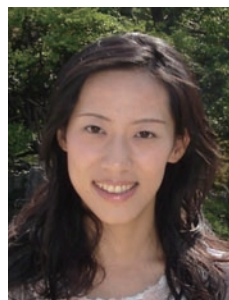
He received the B.A., M.A., and Ph.D. in psychophysics from the University of Tokyo in 1987, 1989, and 2000, respectively. He joined NTT in 1989. From 1992 to 1993, he was a Visiting Scientist at the University of Wisconsin (Prof. Richard Warren's laboratory), USA. Currently, he is a Visiting Professor in the Department of Information Processing, Tokyo Institute of Technology (2006–), and PI (principal investigator) of a JST CREST project on implicit interpersonal information (2009–). He has been investigating functional and neural mechanisms of human cognition, especially auditory perception, cross-modal and sensorimotor interaction, and interpersonal communication through the use of psychophysical experiments, neuroimaging, physiological recordings, and computational modeling.



### Makoto Yoneya

Researcher, Sensory Resonance Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories.

He received the B.E. and M.Sc. in engineering from the University of Tokyo in 2010 and 2012, respectively. He joined NTT Communication Science Laboratories in 2012 and has been studying biological signal processing, especially of eye movements. He is also interested in decoding people's thoughts based on brain or neural activity using machine learning methods and has researched decoding of 'internal voice' using magnetoencephalography signals and multi-class SVM (support vector machine). He is also studying auditory signal processing and is currently developing a mathematical model of auditory salience. He received the 2011 Best Presentation Award from the Vision Society of Japan. He is a member of the Acoustical Society of Japan (ASJ), the Japan Neuroscience Society (JNSS), and the Association for Research in Otolaryngology (ARO).



### Hsin-I Liao

Research Associate, Sensory Resonance Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories.

She received the B.S. and Ph.D. in psychology from National Taiwan University in 2002 and 2009, respectively. She joined NTT Communication Science Laboratories in 2012 and has been studying auditory salience, music preference, and preference of visual images. She has also explored the use of pupillary response recording to correlate human cognitive functions such as auditory salience and preference decision. During 2007–2008, she was a visiting student at California Institute of Technology, USA, where she studied visual preference using recorded eye movements and visual awareness using transcranial magnetic stimulation. She received a Best Student Poster Prize of the Asia-Pacific Conference on Vision (APCV) in 2008, a Travel Award of the Association for the Scientific Study of Consciousness (ASSC) in 2011, and a Registration Fee Exemption Award of the International Multisensory Research Forum (IMRF) in 2011. She is a member of the Vision Sciences Society (VSS) and JNSS.



### Shigeto Furukawa

Senior Research Scientist, Supervisor, Group Leader of Sensory Resonance Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories.

He received the B.E. and M.E. in environmental and sanitary engineering from Kyoto University in 1991 and 1993, respectively, and a Ph.D. in auditory perception from University of Cambridge, UK, in 1996. He conducted postdoctoral studies in the USA between 1996 and 2001. As a postdoctoral associate at Kresge Hearing Research Institute at the University of Michigan, USA, he conducted electrophysiological studies on sound localization, specifically the representation of auditory space in the auditory cortex. He joined NTT Communication Science Laboratories in 2001. Since then, he has been involved in studies on auditory-space representation in the brainstem, assessing basic hearing functions, and the salience of auditory objects or events. In addition, as the group leader of the Sensory Resonance Research Group, he is managing various projects exploring mechanisms that underlie explicit and implicit communication between individuals. He is a member of the Acoustical Society of America, ASJ (member of the Executive Council), ARO, and JNSS.