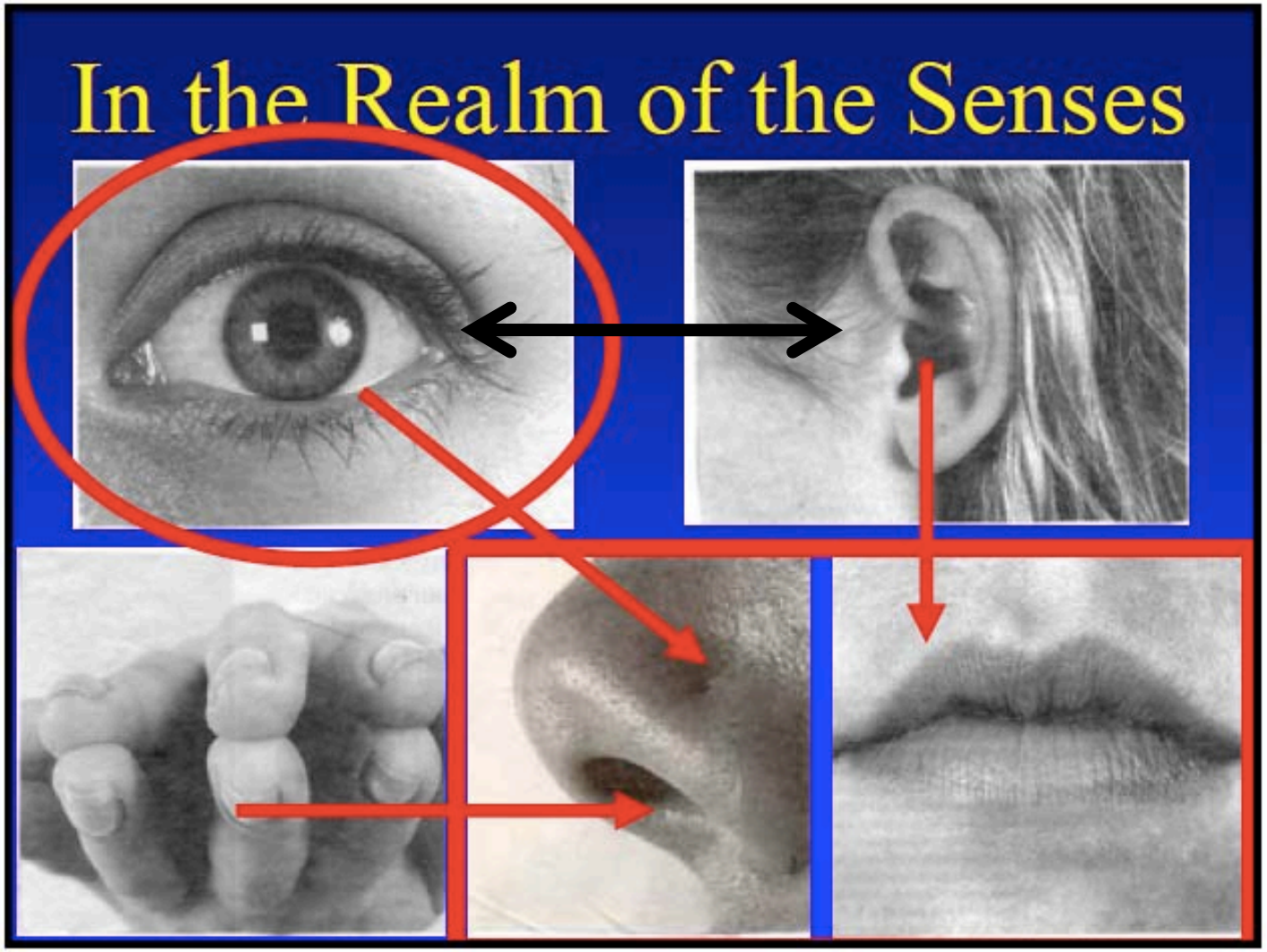
Sensory substitution, and the third kind of "*qualia**"

Shinsuke Shimojo, Noelle R. B. Stiles
& Vikram Chib

Division of Biology / Computation & Neural Systems,
California Institute of Technology

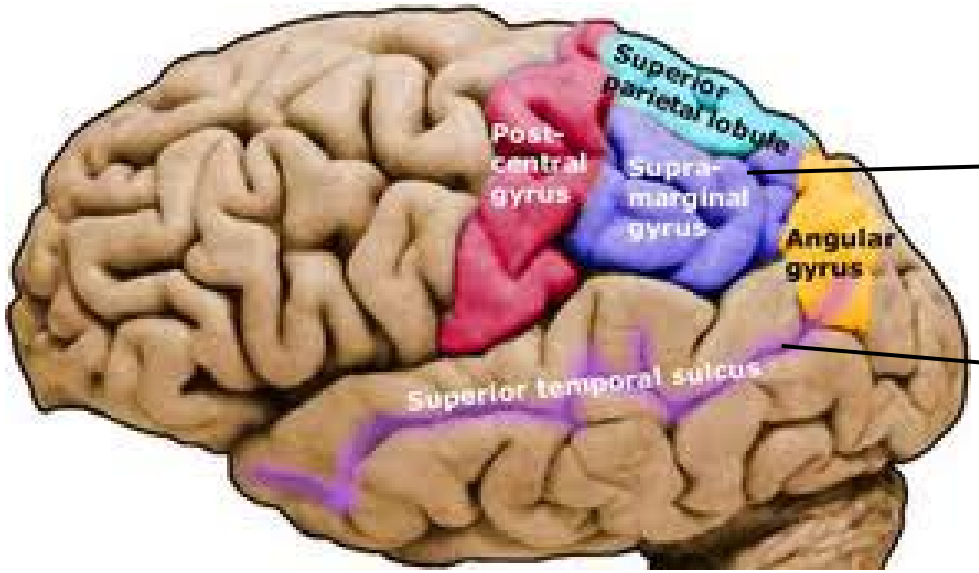
Supported by NSF.GRFP, JST.CREST, Della Martin Fund

* Subjective sensory quality unique to modality;
Nothing to do with "qualia as a hard problem."



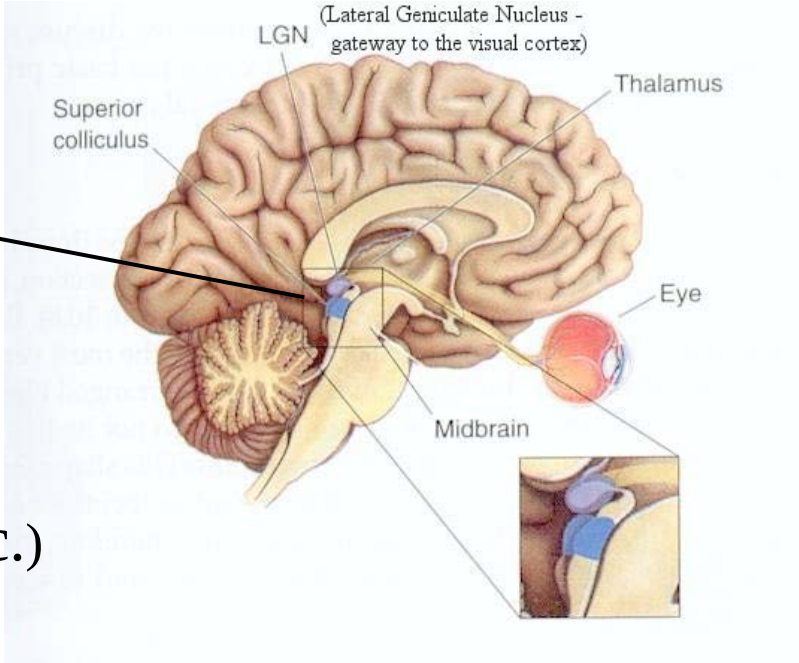
By C. Spence

Neural correlates of multisensory processing



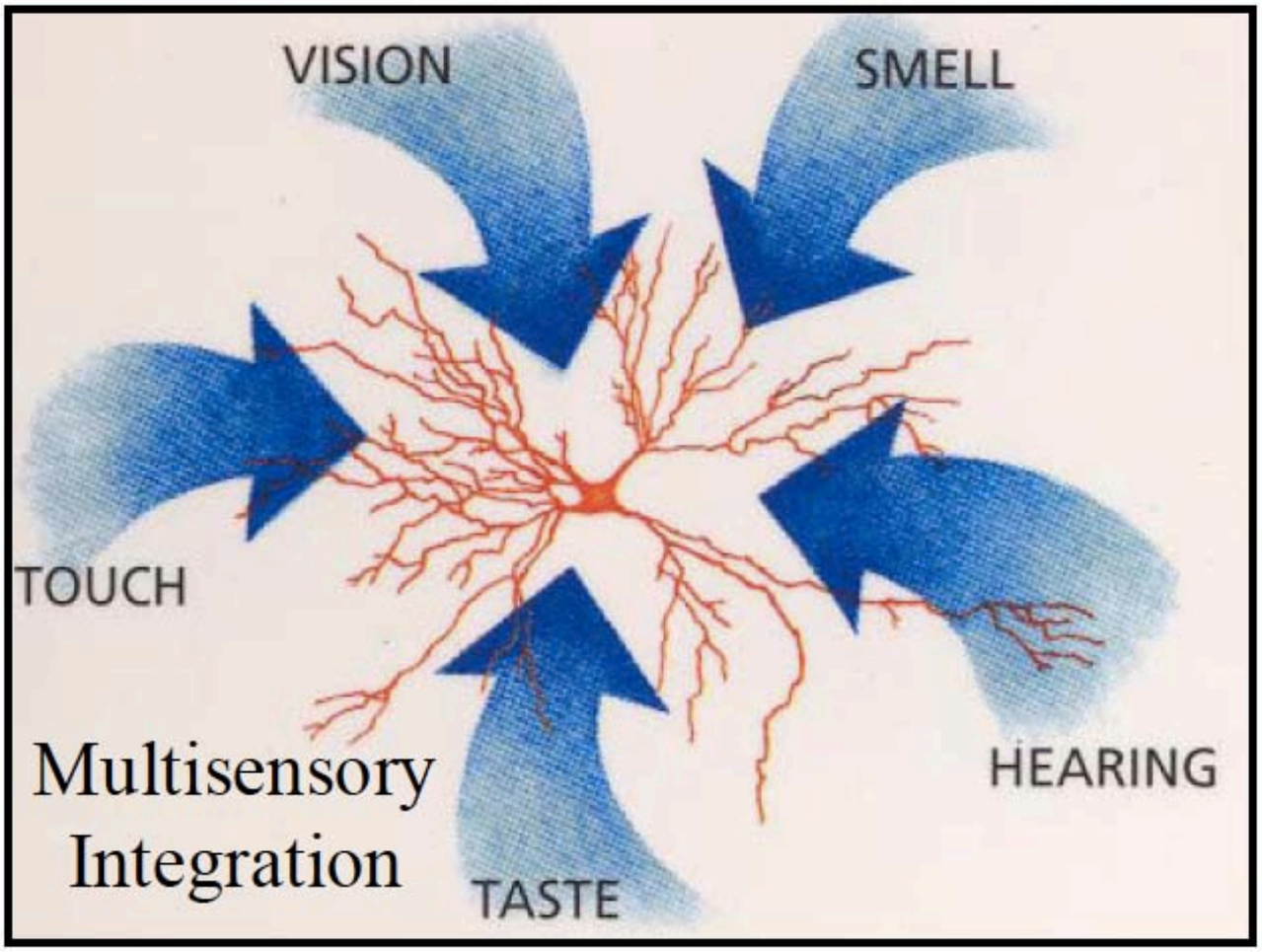
Parietal regions

STS
(Superior Temporal Sulcus)



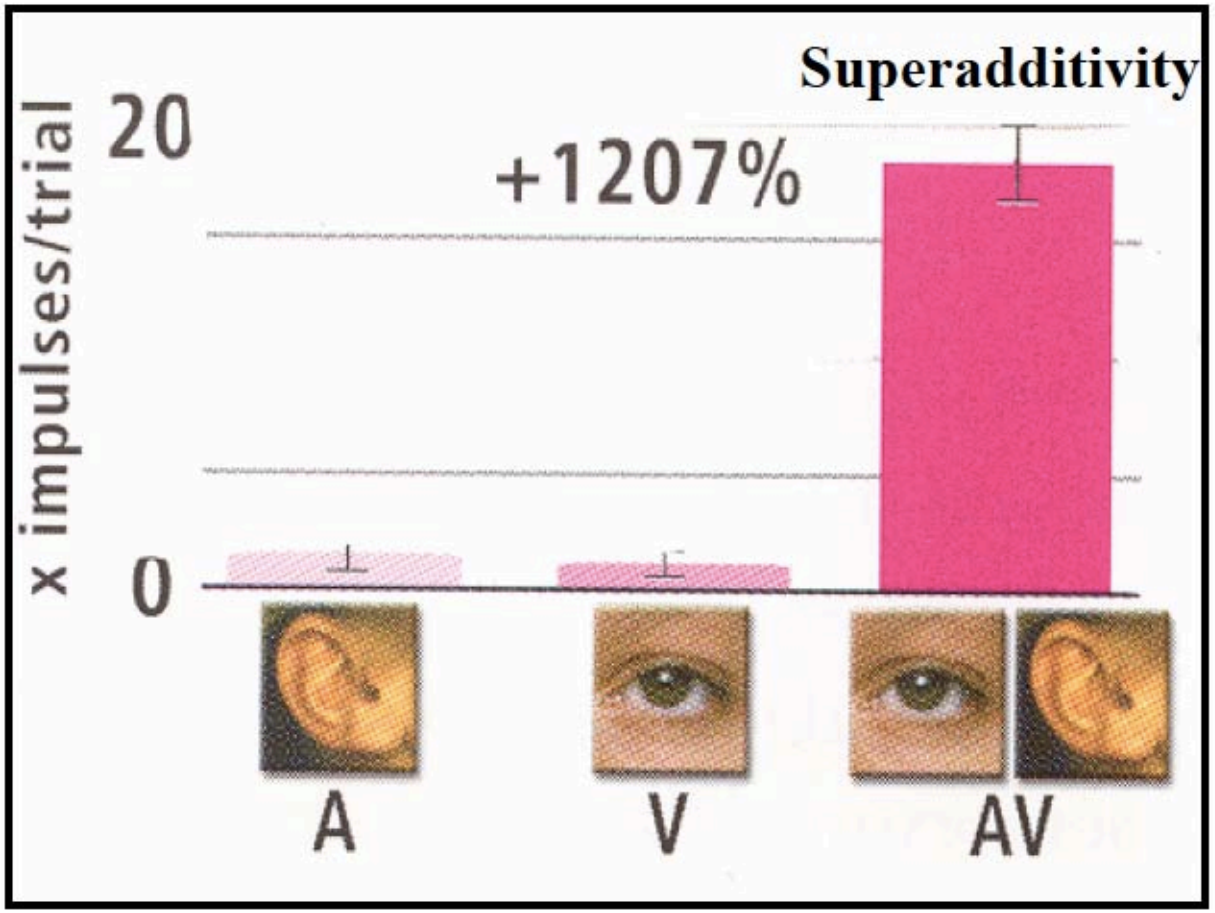
SC (Superior Colliculus)

Primary sensory cortices
(Visual, auditory, somatosensory, etc.)



* Sensory integration at the neuronal level.

* Characteristics of multisensory neurons in the SC (B. Stein)



Firing rates of the SC neurons : AV >>> A = V

By C. Spence

Double Flash Demo.

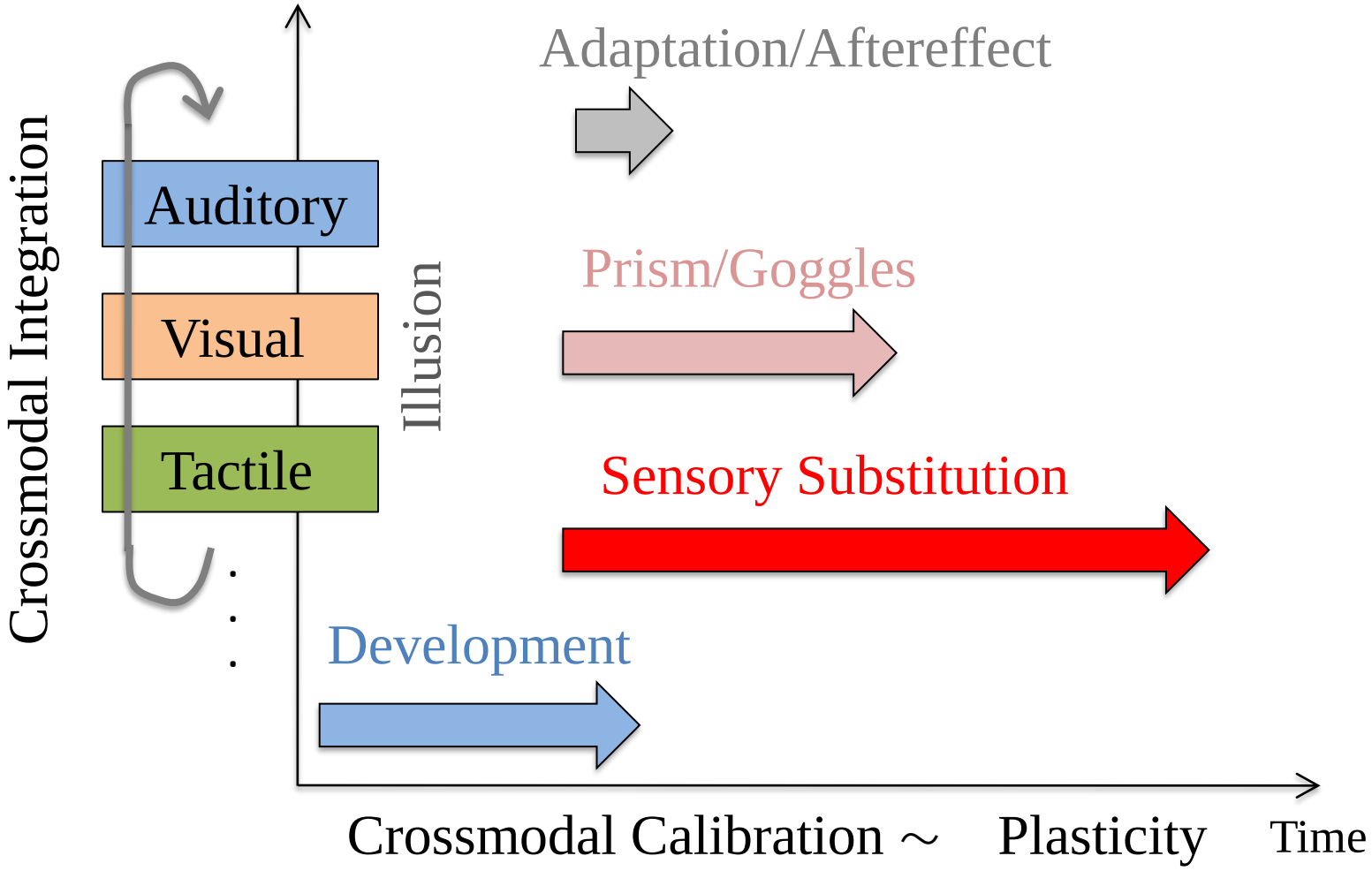
1 flash, 1 sound

1 flash, 2 sounds

How many flashes are there?

(Shams et al., *Nat.*, '99)

Crossmodal - various approaches



Sensory Substitution (SS)

* Translate information from one modality to another
* First device was a back stimulator which translated vision-to-somatosensation (Bach-y-Rita 1969)
* The newest generation is a tongue stimulator (the BrainPort).
* Both V-T type. V-A (Vision to Audition) devices also exist.

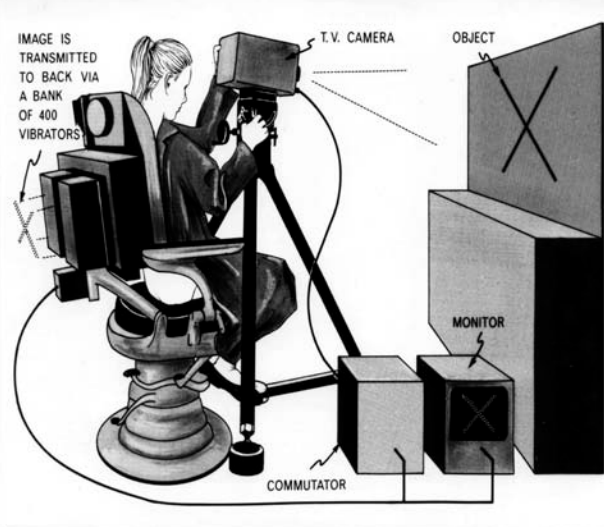


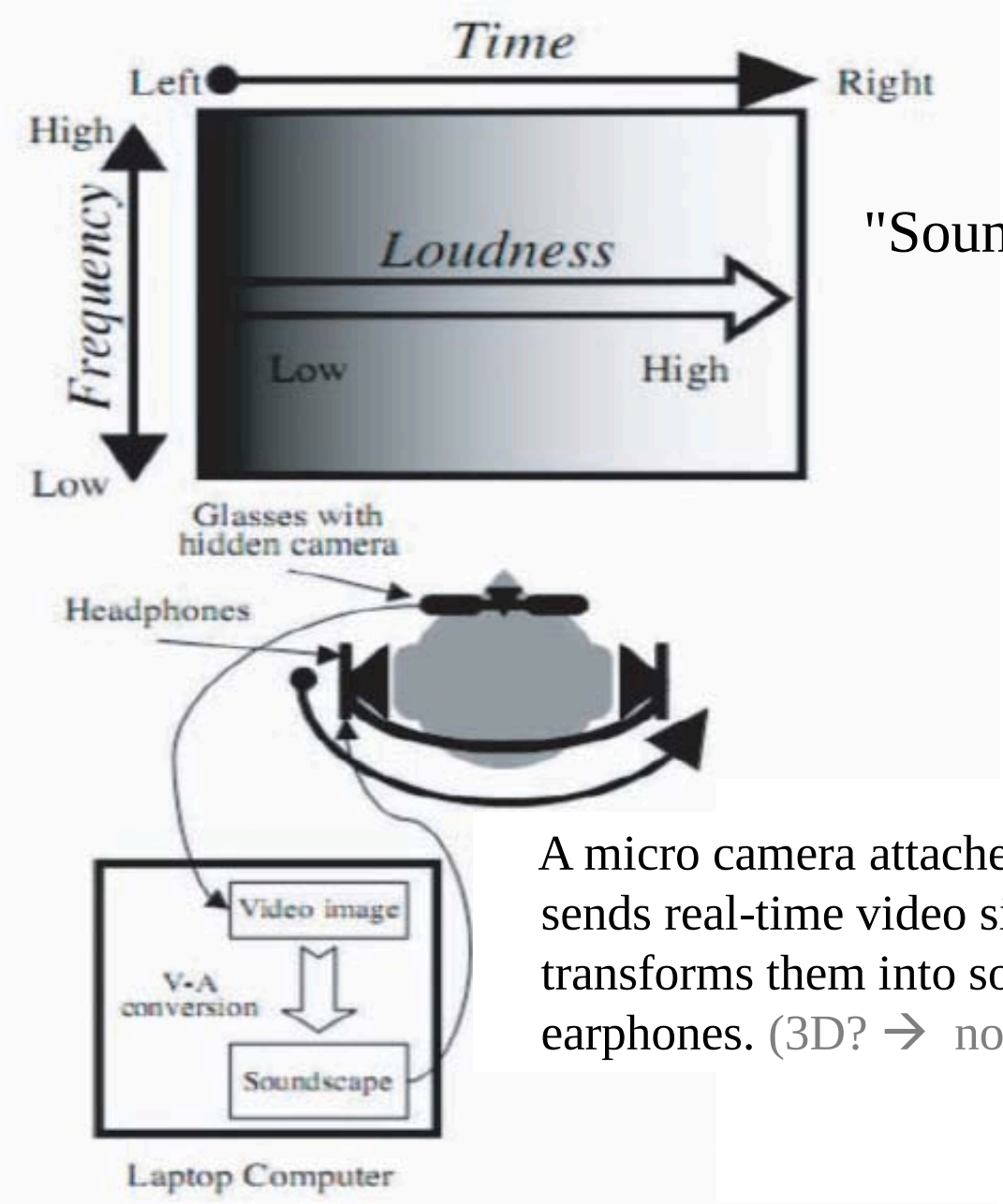
Image of the Brainport prototype

Vision to Audition (V-A) SS

- Several devices exist with different V-to-A encoding methods, w. different ***mapping principles***
- CASBLIP: Cognitive Aid System for Blind People
  - five co-linear fixed points in 3D sound to identify obstacles
- PVSA: Prosthesis for Substitution of Vision by Audition
  - Each pixel in an image is assigned a sound frequency range
- **The vOICe**
  - The bottom to top pixels of a pixel column are assigned frequency ranges from low (bottom) to high (top)
  - The column scans across the image (typically at one hertz)
  - Brightness is translated into volume

→ ***None commercially available, device development still ongoing (w. 40-45 millions market).***

- The **vOICe**, Peter Meijer, 1990s



"Sound scope"

A micro camera attached to the sunglasses. It sends real-time video signals to a mobile PC. It transforms them into sounds, and send out to earphones. (3D? → not yet.)

The vOICe

- Invented by Peter Meijer in 1992 (details at www.seeingwithsound.com) (Meijer 1992)
- Some late-blind users claim that they have the "experience" of vision when using the vOICe
- Several studies have shown lower-level visual activation (BA 19, BA 18, BA 17) when blindfolded sighted and early blind subjects subjects used auditory or tactile sensory substitution devices (Poirier, 2007)
- Higher level visual cortices (in particular LOtv) are activated when recognizing object shape with the vOICe (Amedi, 2007)
- Case study (late blind, N=1): TMS deactivation of regions of occipital peristriate cortex impairs a vOICe user's ability to recognize objects (Merabet, 2009)

Image of a vOICe prototype

The Sound of the vOICe

A horizontal line: the image (above) and the sound (below)

A vertical line: the image (above) and the sound (below)

A slanted line: the image (above) and the sound (below)

Current Focus

Goal of SS: to give blind people "vision."

But what is it like to have "vision"?

→ Seemingly obvious, but not really.

(cf. "What is it like to be a bat?" Nagel, T. '74)

→ ***Is seeing merely visual?*** (L. Albertazzi)

Noelle R. B. Stiles

Vikram Chib

"Vision-like" processing (as opposed to auditory) - Evidence?

1. Phenomenological
2. Neural (fMRI)
3. Functional, psychophysical
4. New "twists"

Note: The data presented here will all be with sighted.

1. Phenomenological(, and daily behavioral)

vOICe superuser, PF

"PF was born in 1956 and became blind ***at the age of 21*** as a result of an industrial accident. She currently has ***a small amount of light perception in the left eye only***, but the right eye was burned out entirely. Before using the vOICe she relied on a cane and guide dog.
She came across The vOICe software in 1998 and began ***using it immersively from 2000***. She uses the standard settings of 1 s refresh rate and normal contrast (i.e. bright mapped to loud). She has taken part in several research studies into the vOICe (Amedi et al., '07; Merabet et al., '09)." (Ward & Meijer, '09)

Pat Fletcher's behavior and reports

- Walks down corridor, sees table, sees cup on table
- Bumping into walls in the middle of the night.
- "I'm not aware of the sound, I just see."
- "It's like looking through blurry glasses."

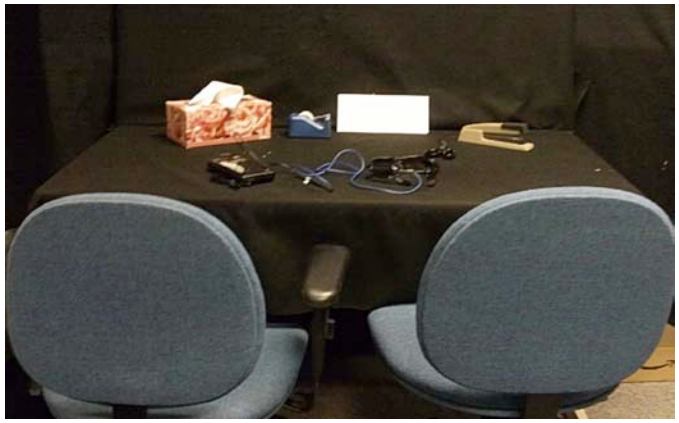
- fMRI tests show *Visual Cortex* responds to soundscapes.

- However, note that she is a late blind (lost sight at age 21)
- Took *3 months* of daily use to learn.
- Still uses it for hours every day (>1 decade).

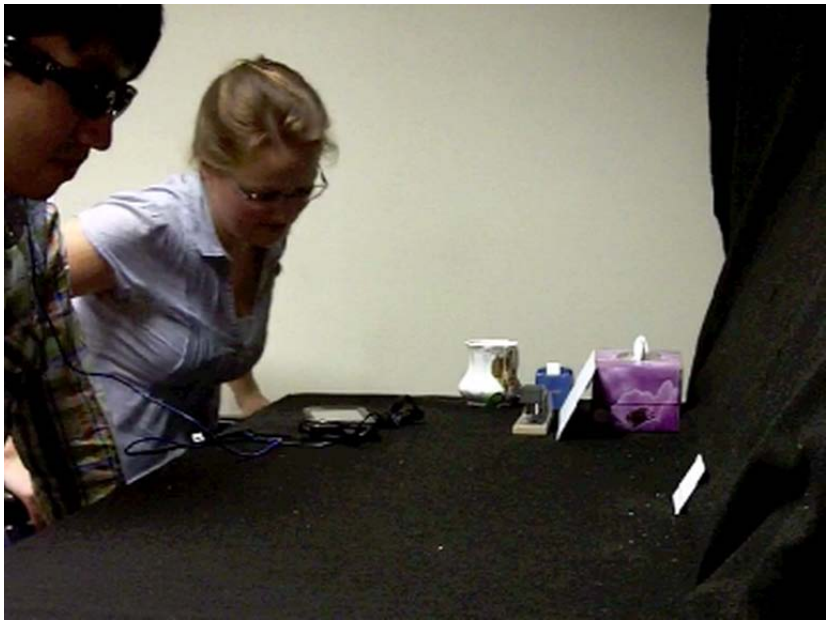
vOICe Training: Recognition Performance

- Five typical office objects are introduced to the subject (**a tissue box, a tape dispenser, an envelope, scissors and a stapler**)
- Subject (sighted, blindfolded) uses the vOICe to locate and identify the object in front.
- The accuracy of the last ten trials is recorded

→ 90～100% correct identifying familiar objects, by 2～6 hrs. of training.



Assessment setup. The subject sits in chair and identifies the object that is placed in in front.



1. Phenomenological(, and daily behavioral) - Summary

1. "Vision-like" phenomenological experiences, at least in some late-blind, super-users.

2. Can discriminate familiar objects, after several hrs of vOICe training. (Head movements useful?)

3. However, it might have been accomplished ***by executive control and cognitive strategy?*** (if so, *different from "seeing"*)

"Vision-like"? --- Automatic & Effortless

2. Neural (fMRI)

* Activation of the ***"where" visual pathway*** (for object localization) from ***vOICe auditory stimuli ?***

* ***Mapping from visual field to visual cortex activation*** (early?) with vOICe ?

fMRI Exp. Procedures

Pre-training fMRI experiment

vOICe Training

Post-training fMRI experiment

- Pointing task w. vOICe training (1 wk, 5 hrs. total)
- fMRI before & after.

(Unpublished data)

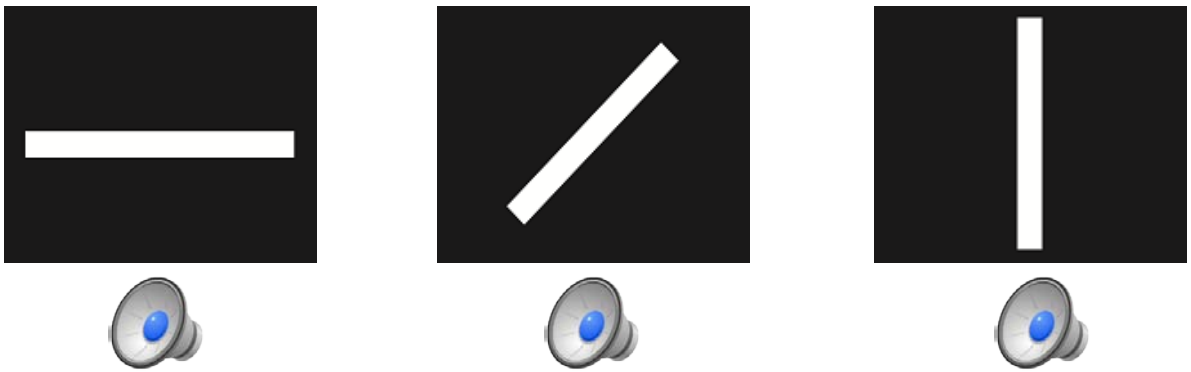
Perceptual constancy

* Retrieving features of of not proximal(retinal), but rather distal stimulus(=object).

* size (distance), shape, brightness, color, depth, etc.

→ Such constancy is a necessary condition for vision-style processing.

* Orientation constancy



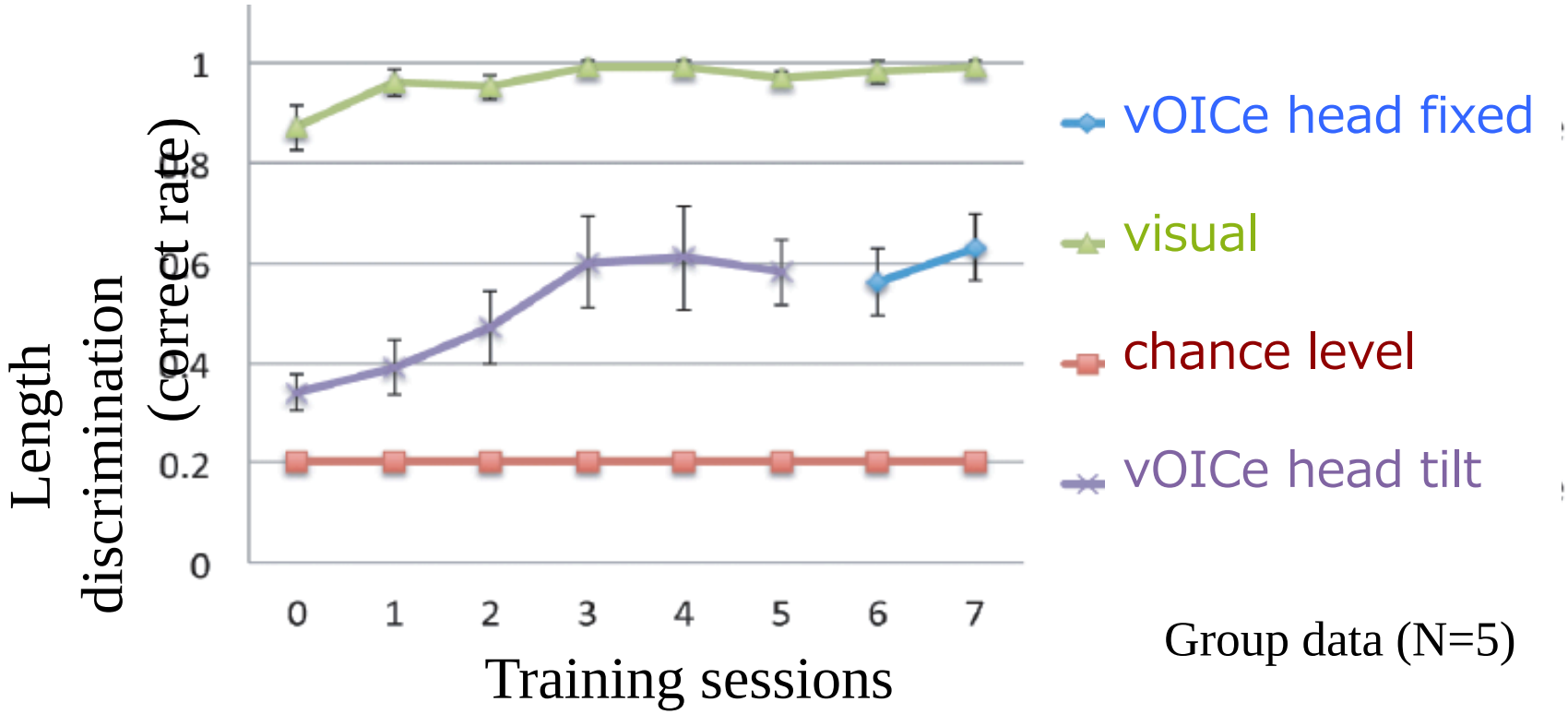
* We know (effortlessly) that these are the same bar.

* With head tilt, retinal image rotates, but external objects unchanged.

Constancy-related task (2)

Length discrimination independent of head tilt

Performance improved w. head tilt movements encouraged, and the learning transferred to the head fixed condition.



Group data (N=5)

3. Functional, psychophysical - Summary

1. Localization of target, discrimination of orientation/length can be improved substantially by 5-10 hrs. of training. ***Perceptual constancy*** had been (partly) established.

2. Performance improved ***better with head free tilt***, as opposed to head fixed.

← Mysterious from the computational viewpoint?
But a matter of course from "affordance" (J. J. Gibson).

NOTE: Congenital/late blind/sighted. The differences are large.
* Late blind > congenital in improvement by training, but
* More plasticity on sensory cortices in the congenital.

Hidden assumption: (Other than by associative learning) there is no intrinsic correspondence among sensory modalities.

eg. Visual stimuli ←→ sounds in the natural world

Maybe these were learned, however …

The latest crossmodal "twists"

So far, we have focused on the Pre-/Post-testing paradigm (i.e. before & after the intensive vOICe training), ***under the assumption*** that un-trained observers cannot do anything with the devise.
It turned out to be wrong!

1) Synesthetic (intrinsic) crossmodal mapping

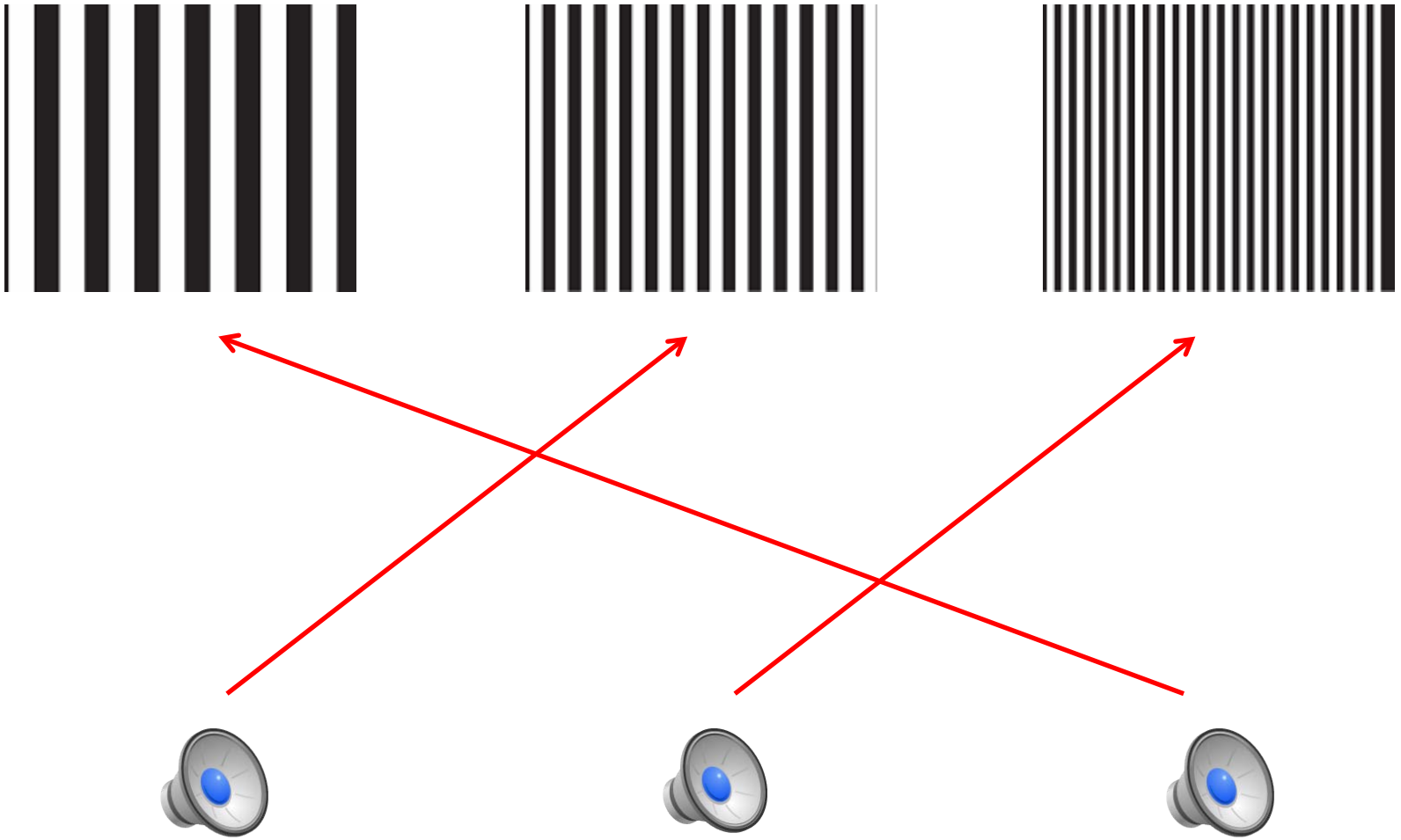
2) Intrinsic, crossmodal mapping
   Sizes of sound/visual object
   Up/down of sound/space
   (Maeda, Kanai & Shimojo, 2003)

→ Could Sensory Substitution (devise development, training) start here, where its ***automatic & effortless*** !?

Testing Synesthesia-like Comparisons with Texture

- Texture is intuitive with vOICe
  - Size in texture: slower change in sound = larger objects
    - Distinguishing among vertical grating of different sizes
    - Distinguishing among randomly placed circles of different sizes
  - Structure in texture: constant frequency = flat line, increasing or decreasing frequency = slanted line
    - Distinguishing among randomly place circles, triangles and squares of similar size
    - Distinguishing among natural textures
- Testing Paradigm
  - Subjects listen to vOICe encoded sound
  - Then choose which image of three presented the sound is "most like" (no training), in a 3AFC task, ***w.o. knowing the mapping principle***

Texture discrimination : Demo.



(Unpublished data)

4. New "twists" - Summary

1. There are "synesthetic" ***inherent crossmodal mappings***, which enable an observer to match soundscapes with visual images, without any training (or even knowledge). (eg. texture)

2. Such inherent mappings should be fully utilized as effective bases of training.

3. It casts a doubt on the conventional visual "primitives."

1) Is the V-A mapping principle ***optimized*** for SS?

2) How to make it ***automatic & effortless***?

→ The texture findings may rescue us from these Qs.

cf. J.J. Gibson "higher-order invariance"
A. Pascual-Leone "Vis. cortex is not just for vision"

4. Blind participants showed a similar intrinsic mapping (A-T).

5. ***A-V connectivity & its change*** are the key.

"Vision-like" processing (as opposed to auditory) - Evidence?

1. Phenomenological
2. Neural (fMRI)
3. Functional-behavioral

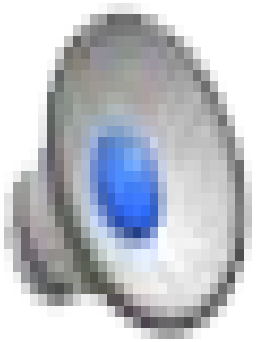
- Vision-like, but ***not quite*** "vision."
- Conscious access to perceptual contents, not only ***after***, but also ***before*** the constancy established.

One can pay attention to quality of sound inputs themselves.
(cf. J.J. Gibson's distinction between "*visual world*" and "*visual field*.")

- ***Mode*** of perception

→ ***The third kind of qualia !***

Neil Harbisson (TED) - an artist with achromatopsia



* Misinterpret nat. sounds as colors, generating vis. artwork (a new synaesthesia?)
* Still "hear" colors; aesthetics follows that of audition
* Qualified for the "***third kind of qualia***" ?
* Related to the intrinsic A-V mapping?

(1) Adding a new associative sensory dimension? Multi-sensory enrichment.

(2) Are colors really perceived? - Y & N. "The third type of qualia."

(3) Is this case special? Or generally applicable to congenital blind, SS, and sighted?

Sensory Substitution and Aesthetics

- Sensory substitution is a multimodal experience (audition transduces, vision interprets)
  - May have aesthetic principles of vision or of audition
  - Aesthetic principles may follow the mode of perceptual experience
    - Late blind => Visual experience
    - Sighted => Auditory experience

- When corresponding image and sensory substitution sound are shown together, there is a ***unique multimodal experience***
  - May generate a unique multimodal aesthetic and new aesthetic principles
  - May amplify aesthetics of both sound and image when displayed together

(Unpublished data)

Summary (and new questions)

1. How would it be like "to see"?

Sensory Substitution not only provides seeds for medical engineering/clinical applications, but also raises more basic questions (*eg.* crossmodal plasticity in the brain).

2. What are "visual primitives"?

Perhaps not static, geometric elements, but rather more dynamic,
active and crossmodal.

3. The third kind of *qualia*?

This is what SS aims for. The "absolute quality" of sensory experience cannot be detached from adaptive behavior.

4. Possibilities for arts?

Adding new sensory dimensions.

END