Philosophy Study

Volume 2, Number 6, June 2012 (Serial Number 11)

DAVID PUBLISHING

David Publishing Company
www.davidpublishing.com

Publication Information:
Philosophy Study is published monthly in print (ISSN 2159-5313) and online (ISSN 2159-5321) by David Publishing Company located at 9460 Telstar Ave Suite 5, EL Monte, CA 91731, USA.

Aims and Scope:
Philosophy Study, a professional academic journal, commits itself to promoting the academic communication about developments in philosophy, covers all sorts of research on Epistemology, Ethics, History of Philosophy, Philosophy of Science, Philosophy of Language, Philosophy of Religion, Philosophy of Mind, Political Philosophy, and other relevant areas, and tries to provide a platform for scholars worldwide to exchange their latest findings.

Manuscripts and correspondence are invited for publication. You can submit your papers via Web Submission system, or e-mail to philosophy@davidpublishing.com (org) or philostudy@yahoo.com. Submission guidelines and Web Submission system are available at http://www.davidpublishing.com.

Editorial Office:

Tel: 1-323-984-7526; Fax: 1-323-984-7374
E-mail: philosophy@davidpublishing.com (org), philostudy@yahoo.com

Abstracted / Indexed in:
The Philosopher's Index, USA
Database of EBSCO, Massachusetts, USA
Chinese Database of CEPS, American Federal Computer Library Center (OCLC), USA
Chinese Scientific Journals Database, VIP Corporation, Chongqing, P.R. China
Ulrich's Periodicals Directory
Cambridge Scientific Abstracts, CSA
Summon Serials Solutions

Subscription Information:
$450 (print/year); $320 (online/year); $600 (print and online/year)

David Publishing Company
9460 Telstar Ave Suite 5, EL Monte, CA 91731, USA
Tel: 1-323-984-7526; Fax: 1-323-984-7374
E-mail: order@davidpublishing.com, shelly@davidpublishing.com

DAVID PUBLISHING

David Publishing Company
www.davidpublishing.com

Philosophy Study

Volume 2, Number 6, June 2012 (Serial Number 11)

Contents

Philosophy Study, ISSN 2159-5313
June 2012, Vol. 2, No. 6, 377-383

Immanent Transcendence in Chinese and Western Process Thinking*

Jan B. F. N. Engberts
University of Groningen

The main classical schools of Chinese philosophy can be characterized as process thinking and their concepts are related to those of Whitehead's philosophy of organism. (Neo-)Confucianism, Daoism, and Chinese Buddhism all find their roots in the *Yi Jing* with a strong emphasis on human experience of the flux of time. These Eastern and Western world views allow attempts to consciously experience experiential reality in terms of immanent transcendence. These attempts are discussed also by considering the functioning of the human brain, emphasizing the importance of brain asymmetry.

Keywords: immanent transcendence, process thinking, *Yi Jing*, Alfred N. Whitehead, neuronal networks, brain asymmetry

1. The Philosophical Tradition in China

In this brief study an attempt is made to provide further evidence for the notion that the fundamental concepts of the major schools of classical Chinese philosophy bear a remarkable similarity to Whitehead's process thinking. With this starting point, our attention will be focused on immanent transcendence, a concept that appears to imply two mutually exclusive issues. However, in both world views, the natural order that human beings experience in their acts of mental perceptions can be understood as equivalent to the general order as enjoyed in their concrete experiences of the dynamic processes in the world around them. In a memorandum to the emperor, written by the Neo-Confucian sage Cheng Hao (1032-85), the first sentence reads: "The laws introduced by the wise monarchs were all based on human nature and in accord with the order in the nature around us." And in the words of Whitehead: "the first determination of order—that is of real potentiality—arises out of the general character of the world" (1929, 103).

It will be argued that human experiences of the essential processes in the world around us can be resolved in patterns of brain functioning of the specific individuals, and ultimately involving activities of both hemispheres of the brain.

The three major philosophical worldviews in China, Daoism, (Neo-)Confucianism, and Chinese Buddhism (all with strong roots in the *Yi Jing*) all can be characterized as process thinking (Engberts 2010). That means

* **Acknowledgments**: I like to express my gratitude to Ron Phipps for many years of inspiring discussion of process thought. I thank Dr. Qian Li and Mrs. J. J. Dong for stimulating discussions of Chinese philosophy and for putting the Chinese characters into the text.

Jan B. F. N. Engberts, Ph.D., Emeritus Professor of Physical Organic Chemistry, Stratingh Institute of Chemistry, University of Groningen, The Netherlands; main research fields: Organic Chemistry in Water, Amphiphilic Molecules, Synthesis and Applications, Western and Chinese Process Philosophy. E-mail: J.B.F.N.Engberts@gmail.com.

that natural existence consists in, and is best understood in terms of, *processes* rather than *things*, in modes of change rather than fixed stabilities. The universe is comprised of a coordinated group of dynamic changes in the complexity of reality, an organized family of occurrences that are systematically linked to one another either causally or functionally (Rescher 1996). Chinese process thought, starting with the *Yi Jing* (易经, *Book of Changes*), is pre-scientific and rests largely on comprehensive observations and intuitive notions, based on specific ultimate categories that were layed down later in history in sacred books and which were considered the heritage of great sages of antiquity (Cheng 2011a). These notions include *yin* (阴) and *yang* (阳) in the *Yi Jing*, *dao* (道) in Daoism, *ren* (仁) in Confucianism, *li* (理) and *qi* (气) in Neo-Confucianism, and *wuzhi* (无执, Skrt.: śūnya) in Chinese Buddhism, although several other categories compete in importance over the ages during which classical Chinese thought has been created. In the words of Arthur F. Wright, “the Chinese culture rests on an indigenous body of thought, remarkable for its perennial vitality and for its intimate relation with every Chinese value and institution” (1953, 1-18).

The above categories need some more detailed attention. *Yin* and *yang* are symbols for physical and mental activities, which are ultimately combined in systems of *gua*’s (trigrams, hexagrams) that can be viewed as symbolic patterns of change. *Yang* is represented by a solid line and *yin* by a broken line. *Yin* is quiescence, *yang* is activity; *yin* is soft, *yang is hard*; *yin* is female, *yang* is male; *yin* is the moon, *yang* is the sun; *yin* is water, *yang* is fire; *yin* is docile, *yang* is dominant. *Yin* and *yang* are competitive and complementary forces or streams of energy. In later times, particularly in the Song dynasty (960-1279), the Neo-Confucianists emphasized that the alternation of *yin* and *yang* generates the five phases: water, fire, wood, metal, and earth. These phases should not be interpreted as substances but rather as metaphors for activities, water for what goes down, fire for upward movements, wood for what can be bowed or straightened, metal for what adapts itself through changes, earth for what can let seeds grow. From the five phases, the “myriad creatures and things” come into existence.

Nothing is completely *yin* or completely *yang*. *Yin* and *yang* potentialities are realized in time and in change in the system of the *gua*’s. *Yin* and *yang* are functioning in a single, harmonic, and cosmic hierarchy, in the experiences of the human being as well as in the processes occurring in nature. This way of thinking has been called “correlative.” According to Richard J. Smith, “conceptions are not subsumed under one another but placed side by side in a pattern” (2008, 32); and “things behave in certain ways not necessarily because of prior actions or impulses of other things but because they resonate with other entities and forces in a complex network of associations and correspondences, conditioned by variables of time, space and patterns of movement” (2008, 32). These basic notions, as first laid down about 3,000 years ago in the *Yi Jing*, have dominantly influenced ancient, medieval, and modern Chinese culture, ranging from the realms of language, literature, art, calligraphy, music to philosophy, politics, law, military affairs, social life, medicine, and science (Smith 2008). An interesting example in modern science involves a study of the molecular structure of the CPAF (chlamydial protease-like activity factor) protein which was shown to be a homodimer in which two identical catalytic domains were characterized as the *yin* and *yang* of the protein structure (Huang, et al. 2008; Qiu 2009).

The deep significance of the foundations developed in the *Yi Jing* have been accepted in Daoism and (Neo-)Confucianism and have also direct relevance for Chinese Buddhism. *Dao*, the way, conceived as the dynamics seen in “all under heaven” (Ames and Hall 2003), can be described in terms of the *dao* of heaven, the *dao* of man, and the *dao* of earth, which all reflect the fundamental order of the Universe.

Ren, an ultimate category in Confucianism, is the inner moral principle, viewed as the internalization of the Way of Heaven in all human beings. It stands for the life-giving creativity and involves a loving relation between you-and-me, humanity, goodness, and benevolence (Wang 2008). In *The Great Learning* (大学), the first of the four books of Confucianism that formulate the basic teaching of the later Confucian tradition, the first line reads: “The way of great learning lies in letting one’s inborn luminous virtue shine forth, in renewing the people and in coming to rest in perfect goodness” (Gardner 2007, 3).

Three to four thousand years ago, a world view emerged in China in which the human being was considered as part of the nature surrounding him. The recurring natural disasters, including the enormous floods that costed the lives of numerous people and that destroyed the harvests of the farmers, made it clear that the people and his environment formed an unbreakable unity. This conception would never leave the Chinese people. One has to live in harmony with the dynamics and cosmic rhythms of the universe. These ideas have been implanted long ago in the ancient folk religion and philosophy *Feng Shui* (风水, “wind and water”), often with anthropomorphic images of deities and later containing Confucian, Daoist, and Buddhist view points.

The traditional Chinese philosophies are all based on the same concept as often expressed in metaphors. For example, Kongzi (Confucius), looking at a river, said, “passing on like this, it never ceases, night and day. Nature and the human world should be conceived as participating in a creative harmony” (Lau 1979, 98).

The intuitive feeling that all humans are connected to each other has recently been verified by neuroscientific discoveries. The basis of empathy is the notion that we are “moved by” the conscious observation of other humans, as affected by our mirror neurons that are activated by our direct experience of, for example, the suffering of other people (Iacoboni 2008). These mirror neurons are directly connected with our emotional brain centers and also with our pre-SMA, a brain center important for complex motor sequences and motor planning. In other words, empathy and benevolence involve basic and natural forms of conciousness which find their basis in the functioning of the human brain.

Important metaphysical concepts, largely developed by Zhu Xi (朱熹 1130-1200), include the categories *li* and *qi*. *Li* is the heavenly principle (*tian-li*, 天理) of permanence amid flux, the ideal pole, and “above shapes.” It is involved in perpetual change and becoming through the two modes of *yin* and *yang*. Zhu Xi spoke of *tian-li* as an all-embracing, transcendent absolute that is immanent in human minds and hearts (Ching 2000). *Qi* is the material principle, giving shape and actuality to *li* and causing creativity to be immersed in the process of concrescence.

In the different schools of Chinese Buddhism there is a strong focus on *śūnya* (Skrt.: Pāli, Cheng 2011b), often interpreted in the Chinese scripts as *wuzhi* (无执, no clinging). Equally important is the notion of compassion (*Karunā*, Skrt.: Pāli), the greatest virtue of a bodhisattva. The Buddha even spoke about *Mahakáruna*, great compassion, all human activities aimed at reducing the suffering of other living creatures. The meaning of *Karunā* bears a direct resemblance to *ren*, and also reflects the doctrine that all human beings and nature belong to the same comprehensive harmony of the universe. The indivisible unity of human existence and the surrounding world unfolds itself in discovering and exploring of the meaning of life.

2. Whitehead’s Process Thinking

Alfred N. Whitehead’s scientific process philosophy bears a remarkable resemblance to classical Chinese thought (Engberts 2010). In *Process and Reality* (1929), Whitehead (1861-1947) puts forward his Category of the Ultimate. These ultimate notions are *creativity* and *many and one*, which are inexplicable in terms of higher universals. Creativity is that eternal aspect of the universe by reason of which there is an endless becoming of actual entities, i.e., drops of experience, complex and interdependent. It generates the transition of time and the extensive relationships of time. It manifests itself in the becoming of all *actual entities*, the "atomic" facts of life, which are not material substances but activities of realization. According to Whitehead (1929) the notion of change should be defined as the difference between actual occasions comprised in some determinate event. Creativity is the final metaphysical principle that there is an endless passage from disjunction to conjunction, involving the creation of novel entities ("one") other than the entities given in disjunction ("many"). The ultimate reality bestows an orderly character on creativity so that the becoming of *actual entities* is not purely chaotic, but rather functions in the selection of repeatable patterns ("*eternal objects*," pure potentials of the universe) that are enacted in the world as subjective forms of experiences. Concrescence is then defined as a process of acquiring definiteness by a series of decisions in which actual entities are selected or rejected. Therefore, change is the description of the adventure of *eternal objects* in the evolving universe of actual entities. For our discussion it is important to emphasize that Whitehead too argues that the universe and the human world should be conceived as part of an essential oneness.

The paired concepts of *yin* and *yang*, as briefly outlined above, are not inconsistent with Whitehead's process thinking. In fact, the idea of fundamental dualities strongly appeals to him, as pointed out in one of his later books:

> The Universe is dual because, in the fullest sense, it is both transient and eternal … the final actuality is both physical and mental … the Universe is *many* because it is wholly and completely to be analyzed into many final actualities … the Universe is *one* because of its universal immanence … throughout the Universe there reigns the union of opposites which is the ground of dualism. (Whitehead 1933, 190)

Whitehead's thought can be summarized in much greater detail, but the present notions appear sufficient for the point I like to make in this essay.

It will be clear, however, that despite the different cultural matrices, process thinking is an essential concept in both Whitehead's philosophy and in ancient Chinese thinking. This is now well recognized in modern China and led to the establishing of about twenty Centres for Process Studies at different Chinese universities (Yang 2010).

3. Immanent Transcendence in Human Experience

Consciousness presupposes experience, and the question can be asked how our brain allows experience to come into being. In an in-depth study, Kenneth K. Inada (2008) has emphasized the essential nature of *immanent transcendence*, particularly in Asian thought. Transcendence (*chao yue xing*, 超越性) can be viewed as "exceeding or surpassing in degree or excellence," the property of rising out of or above other things. It is the property of being of a higher order and outside the observed material world. Immanence (*nei zai xing*, 内在性) means "existing, operating, or remaining within," it is in itself, encompassing the material world. *Immanent transcendence* is an inherent vision of human experiences that transcends classical cognitivism, that is our direct empirical and rational understanding of occurrences around us. This elemental aspect of experience invokes consciousness and cognition as abstractions and functioning in our daily life. As noted by Whitehead (1929), immanence and transcendence are both characteristics of an actual entity. The immanence of a fundamental order gives reason to believe that pure chaos is intrinsically impossible. And the transcendence involves that the actual occasions are activities of realization, complex combinations of energy events that exhibit an order that we are used to call the laws of nature.

In the Western natural sciences implying a materialistic world view, arguments have been presented for the existence of a universal consciousness, defined as the reprensentation of a quantum knowledge field that encompasses the whole universe. It is then envisaged that the human brain can occasionally operate as a bilaterally operating interface between the individual and universal consciousness (de Quincey 2002; Henry 2005).

4. The Specific Identity of a Human Being

Let us first have a brief look at the functioning of the brain in an attempt to relate the philosophical notions laid down in process thought with aspects of brain functioning. Particular emphasis will be placed on consciousness and on a definition of the human brain. This part of our discussion follows rather closely the views of Susan Greenfield as summarized in her recent book *The Pivate Life of the Brain* (2000). Consciousness, with feelings as its most basic form, can be viewed as a defined level of brain organization, i.e., highly transient groupings of neurons (brain cells) that are in a dynamic process of changes. These dynamic neuron assemblies can accomodate varying degrees of subjective, vanescent consciousness. The size of the neuronal networks appears to determine the depth of consciousness, and depends on a variety of factors, including (1) the number of already existing neuron networks, (2) chemical factors that should function properly (neurotransmitters, (pro)hormones, peptides), (3) the degree of stimulation where strong sensory stimulation activates a greater temporary size of a working neuron assembly, and (4) the speed of neuron recruitment by rival assemblies, that determines the time available for realization of the full potential. It should be noted that neuron activation takes about 0.25 seconds, whereas consciousness of sensory perception needs about 0.5 seconds.

Important for consciousness are both the cortex and the thalamus, the latter located deeply into the brain with extensive connections to outer reaches of the brain. Finally it is important to stress that there are no committed neurons or genes devoted to consciousness.

The human mind can be viewed as a system of networks of neurons, formed after birth and reflecting individual experiences. This is the most plausible physical basis for personalizing the brain into an idiosyncratic conglomeration of experiences, generalizations, prejudices, and propensities. All the time experiences leave their mark and in turn determine how we interpret new experiences (Greenfield 2000). During our whole life, the mind takes shape and the individual mind will color the way one feels at varying times. These extensive numbers of operational neuronal networks change and evolve, but rather slowly over longer time frames. At a particular time, our self lies in the totality of the brain's wiring, called the "connectome" by neuroscientists (Seung 2012). But both our consciousness and our brain appear to exist in a dynamic, kinetic stability which, as argued by Pross (2008), is characteristic for far from equilibrium systems as they do occur in life processes. But it will be clear that the mind is *not* a unique moment of consciousness.

Taking into account the above neurophysiological insights, one can now ask whether the notion of consciousness and mind provides the possibility to define a specific identity of an individual human being. The Buddhist view can be summarized as "continuity, no identity" (van de Beek 2008), in accordance with the dynamic nature of both consciousness and the human mind. The following long quote from Whitehead is fully consistent with this notion:

> The personal identity is the thing which receives all occasions of the man's existence. It is there as a natural matrix for all transitions of life, and is changed and variously figured by the things that enter it; so that it differs in its character at different times. Since it receives all manner of experiences into its own unity, it must itself be bare of all forms. We shall not be far wrong if we describe it as invisible, formless, and all-receptive. It is a locus which persists, and provides an

> emplacement for all occasions of experience. That which happens in it is conditioned by the compulsion of its own past, and by the persuation of its immanent ideals. (1933, 187)

5. Interpretation of Experiental Reality and the Hemispheres of the Human Brain

On the basis of the above discussion, I suggest that it is particularly the human brain that is able to function as the basis for immanent transcendence. The neuronal patterns that have been created and continuously evolve as a response to a specific cultural and physical environment may provide the individual with a world view that is used for ordering of external experiences. For example, in classical Chinese poetry, there are poems in which the emotions of the poet and the world around him are resolved into each other. The consciousness of the poet is submerged in the universe, providing feelings of peace and harmony (Liu 1956).

Cultural variation in neurobiological mechanisms underlying consciousness has already been reviewed (Chiao, Li, and Harada 2008). It seems clear now that cultural differences may find their origin in specific, culture-associated patterns of brain activation.

One can refer to immanent transcendence in a metaphoric sense by seeing it as "turning on the light" of experiential reality in our brain. The light, however, can have different "colors" associated with culture-specific patterns of brain activation. Going a step further, one could assume that these "colors" reflect dominant brain activities originating predominantly either from the right or left hemisphere of the human brain (McGilchrist 2009). Both hemispheres appear to possess different "personalities." The left hemisphere may be best viewed as a serial processor, specialized in linear and methodic thinking. By contrast, the right hemisphere is primarily functioning as a parallel processor and is more strongly involved in feelings of empathy and in intuitive notions. But the two hemispheres are known to communicate intensively with each other primarily via a string of nerves (corpus callosum), and a particular world view will never be determined by conscious reflections nestling in only a single hemisphere.

It should be stressed that the basis of the relational unity underlying any and all cultures should remain maintained regardless of their initial distinctness or separateness (Inada 2008). In the words of Zhu Xi, "The sage feels the influences of the minds of all people in the world" (Chan 1967, 93). However, the power of intuition is always precedent to any act of dichotomy or division.

Definitions of ultimate categories, both in Whitehead's metaphysics and in classical Chinese thought, can be interpreted as attempts to define the adequacy of the human experiences, as emerging from immanent transcendence, for an understanding of the universe. Since the pre-scientific process thinking, as laid down in classical Chinese thought, is largely based on intuitive interpretations of direct and comprehensive observations of the outside world, it might be suggested that these world views are dominantly nestled in the right hemisphere of the brain, which is the hemisphere of questions of "how." In this respect, there is an important difference with many of the dominant materialistic schools of Western philosophy which are primarily preoccupied with questions of "what," i.e., mental activities embodied in the left hemisphere (McGilchrist 2009).

A notable difference can also be observed between the particular languages used by the Chinese process philosophers and by Whitehead. The first have a preference for a poetic, evocative language, often employing metaphors, and with a special attention on human life. By contrast, Whitehead is focusing on the natural world and often writes in the technical language of mathematics and quantum physics with insightful logical deductions.

In deep conscious reflection, however, brain asymmetry does not prevent a final and pure interpretation of experiential reality in terms of ultimate notions given to the human being by the functioning of the *complete* human brain. If this can be realized, "the meditator is imposing back on the outside world a complete understanding, an unconditioned compassion" (Greenfield 2000, 158). Immanent transcendence is then not anymore culture dependent. In a metaphoric sense, "the light that has been turned on" is then colorless, that means it is composed of all wavelengths that can be experienced by the human eye.

Works Cited

Ames, Roger T., and David L. Hall. *Dao De Jing: A Philosophical Translation*. New York: Ballantine Books, 2003.

van de Beek, Peter. Unpublished Lecture. 2008.

Chan, Wing-tsit, trans. *Reflections on Things at Hand, Zhu Xi and Lü Tsu-chien (Chin-ssu lu)*. New-York: Columbia UP, 1967.

Cheng, Chung-Ying. "Interpreting Paradigm of Change in Chinese Philosophy." *Journal of Chinese Philosophy* 38.3 (2011a): 339-67.

---. "A Transformative Conception of Confucian Ethics: The *Yi Jing*, Utility, and Rights." *Journal of Chinese Philosophy* 38.1 (2011b): 7-28.

Chiao, Joan Y., Zhang Li, and Tokiko Harada. "Cultural Neuroscience of Consciousness: From Visual Perception to Self-Awareness." *Journal of Consciousness Studies* 15.10-11 (2008): 58-63.

Ching, Julia. *The Religious Thought of Chu Hsi*. Oxford: Oxford UP, 2000.

Engberts, Jan B. F. N. "The Natural Sciences, Classical Chinese Philosophy, Process Thinking, and Brain Lateralization." *Process Studies Supplement* 16 (2010): 1-39.

Gardner, Daniel K. *The Four Books*. Indianapolis, Cambridge: Hacket, 2007.

Greenfield, Susan. *The Private Life of the Brain*. London: Penguin Books, 2000.

Henry, Richard C. "The Mental Universe." *Nature* 436 (2005): 29.

Huang, Zhiwei, Yingcai Feng, Ding Chen, Xiaojing Wu, Siyang Huang, Xiaojun Wang, Xingguo Xiao, Wenhui Li, Niu Huang, Lichuan Gu, Guangming Zhong, and Jijie Chai. "Structural Basis for Activation and Inhibition of the Secreted Chlamydia Protease CPAF." *Cell Host Microbe* 4.3 (2008): 529-42.

Iacoboni, Marco. *Mirroring People—The New Science* of *How We Connect with Others*. New York: Farrar, Straus and Giroux, 2008.

Inada, Kenneth K. "Immanent Transcendence: The Possibility of an East-West Philosopical Dialogue." *Journal of Chinese Philosophy* 35.3 (2008): 493-510.

Lau, Din C., trans. *Confucius, The Analects*. London: Penguin Books, 1979.

Liu, James J. Y. "Three Worlds in Chinese Poetry." *Journal of Oriental Studies* 3 (1956): 278-90.

McGilchrist, Iain. *The Master and His Emissary*: *The Divided Brain and the Making of the Western World*. New Haven, London: Yale UP, 2009.

Pross, Addy. "How can a Chemical System Act Purposefully? Bridging between Life and Non-life." *Journal of Physical Organic Chemistry* 21.7-8 (2008): 724-30.

---. "Stability in Chemistry and Biology: Life as a Kinetic State of Matter." *Pure and Applied Chemistry* 77.11 (2005): 1905-21.

---. *What is Life? How Chemistry Becomes Biology*. Oxford: Oxford UP, in press.

Qui, Jane. "Sexually Transmitted Diseases: Yin-Yang Action." *Nature China Internet Alert*. Jan. 21 2009.

de Quincey, Christian. *Radical Nature: The Soul of Matter*. 2002. Rochester: Park Street, 2010.

Rescher, Nicholas. *Process Metaphysics*. New York: SUNY, 1996.

Seung, Sebastian. *How the Brain's Wiring Makes Us Who We are*. Houghton Mifflin Harcourt Trade, 2012. Reviewed by Christof Koch. "Neuroscience: The Connected Self." *Nature* 482 (2012): 31.

Smith, Richard J. *Fathoming the Cosmos and Ordering the World: The* Yijing *and Its Evolution in China*. Charlottesville, London: U of Virginia P, 2008.

Wang, Liping. "Transcendence or Immanence? Lévinas, Bergson, and Chinese Thought." *Journal of Chinese Philosophy* 35.suppl.s1 (2008): 89-104.

Whitehead, Alfred N. *Adventures of Ideas*. New York: Free Press; Macmillian Co., 1933.

---. *Process and Reality*. 1st ed. 1929. Corrected Edition. New York: Free Press, 1978.

Wright, Arthur F. "The Chinese Language and Foreign Ideas." *Studies of Chinese Thought*. Chicago: U of Chicago P, 1953. 286-303.

Yang, Fubin. "The Influence of Whitehead's Process Thought on the Chinese Academy." *Process Studies* 39.2 (2010): 342-49.

Philosophy Study, ISSN 2159-5313
June 2012, Vol. 2, No. 6, 384-397

Conceptions of Intuition in Poincaré's Philosophy of Mathematics

Olga Pombo
University of Lisbon

The aim of this paper is to contribute to the identification and characterization of the various types of intuition put forward by Poincaré, taking his texts as a laboratory for looking for what intuition might be. I will stress that these diverse conceptions are mainly formulated in the context of Poincaré's controversies in opposition to logicism, to formalism, and in the context of Poincaré's very peculiar conventionalism. I will try to demonstrate that, in each case, Poincaré comes close to a specific tradition (Kant, of course, but also Leibniz and Peirce).

Keywords: Poincaré, intuition, Poincaré's philosophy of mathematics, logicism, formalism, conventionalism, Kant, Leibniz, Peirce

1. Introduction

We know that intuition is a very ambiguous word, a word which we use to designate a capacity or a faculty which we are not able to describe with sufficient accuracy.

From Plato's *noesis* to Aristotle's *nous*, from Descartes' intellectual intuition to Kant's sensible and pure intuition, from Spinoza's *scientia visionis* to Bergson's vital intuition, from Schelling's metaphysical intuition to Husserl's categorical intuition, the word intuition is used with very different meanings. And the etymological significance of the word as a direct, immediate vision does not help us solve the issue as we still would need to understand what is the nature of such a vision (a quasi-perception, an intellectual vision), what does such a vision gives us to see (a real individual object, an ideal entity, an essence, a concept, a relation), or what is the cognitive value of such a vision (a method for valid knowledge, a doubtful process, a discovery device, a guidance for action).

We also know that intuition plays a major role in Poincaré's philosophy of mathematics. It is so important that Poincaré is usually considered a pre-intuitionist, someone who, before Brouwer, could consider mathematics as a free creation of the human mind.

However, Poincaré uses the concept of intuition in a great variety of meanings. He is aware of the plurivocity of the word. He is aware of the need for distinguishing different types of intuition. And he even gives some precise indications in this respect. That is the case of a celebrated distinction between three types of intuition he formulates in *La valeur de la science*. As he writes:

> Nous avons donc plusieurs sortes d'intuitions; d'abord, l'appel aux sens et à l'imagination; ensuite, la généralisation par induction, calqué pour ainsi dire, sur les procédés des sciences expérimentales; nous avons enfin l'intuition du nombre pur,

---

Olga Pombo, Ph.D., Professor of the Faculty of Sciences, Center for Philosophy of Science, University of Lisbon, Portugal; main research fields: Philosophy of Science, Science and Art.

> celle d'où est sortie le second des axiomes que j'énonçais toute à l'heure et qui peut engendrer le véritable raisonnement mathématique. (1970, 33)

But in the very same text and also in other writings, Poincaré uses the word intuition with different meanings. And, even within what could be called one specific type of intuition—we will see this later—he introduces different formulations pointing to diverse conceptions of intuition. That is why it seems legitimate to state that Poincaré provides neither a clear doctrine nor a systematic use of the concept of intuition. The aim of this paper is precisely to contribute to the identification and characterization of the various types of intuition put forward by Poincaré in the context of his anti-logicism, anti-formalism, and very peculiar conventionalism.

2. Against Logicism

The main opposition concerns the fact that for Poincaré mathematics cannot be reduced to logic and that precisely because intuition plays a significant role. In his text "Les mathématiques et la logique," Poincaré begins by pointing out his complete opposition to such a thesis.

Five arguments are put forward. Let us begin with the first three. The *first* is of factual nature. The logicist program is evaluated by its results, that is, by the contradictory consequences of its own development. As Poincaré declares:

> Malheureusement, ils sont arrives à des résultats contradictoires, c'est ce qu'on appelle les antinomies cantoriennes. (1999, 127)[1]

A similar *second* argument is developed in what concerns historical origins of mathematics. If mathematics could have been made just with logic and by logic, its development would have not taken place.[2] That is, logic guarantees the rigor of mathematics but not its invention and progress.

The *third* is a pedagogical argument. For reasons parallel to those of historical development of mathematics, it is not possible to learn and to teach mathematics without the help of extra-logical elements. As Poincaré states repeatedly, logicist methodology for learning mathematics in which there is no place for any kind of intuition is "contrary to all reasonable psychology."[3] That is, logic guarantees the rigor of mathematics but does not provide comprehension.

Several types of intuition are involved in this process of comprehension. To comprehend, students need to see with the eyes of the body. As Poincaré writes in *Les définitions mathématiques et l'enseignement*:

> Dans les écoles primaires, pour définir une fraction, on découpe une pomme. (1999, 106)

Clearly, Poincaré points here to *sensible intuition*. But that need of a sensitive, direct contact with the material entities corresponds to a very preliminary moment, both at the level of individual comprehension and at the level a collective historical process of mathematical development.[4] Soon, this need of sensible images has to be surmounted. As in the history of mathematics, students have to realize, and even to desire, to rise above that initial, rudimentary, primitive level of comprehension.

> Mais il faudrait leur montrer qu'ils ne comprennent pas ce qu'ils croient comprendre, les amener à se rendre comte de la grossièreté de leur concept primitif, à désirer d'eux-mêmes qu'on l'épure et le dégrossisse. (1999, 107)

We understand well the fundamental reason why, for Poincaré, sensible intuition cannot have any relevant role in mathematics, neither in its learning and comprehension nor in its historical development. And it should be so because, for Poincaré, mathematics is independent of material entities. To exist means to be without contradiction. Now, to comprehend mathematics, Poincaré stresses further, students should fulfill their need of thinking with images:

> Sous chaque mot, ils veulent mettre une *image sensible*. (1999, 105, our emphasis)

But what does Poincaré mean by sensible image ("image sensible")? What kind of image is Poincaré pointing to? Note that for Poincaré it is not just young students who show such necessity of thinking with images. Poincaré extends such a need to all men. As he states:

> Comment trouver un énoncé concis qui satisfasse à la fois aux règles intransigeants de la logique … et à *notre* besoin de penser avec des images. (1999, 114, our emphasis)

Of course, when Poincaré speaks of "our need of thinking with images" he is not talking about images provided by sensible intuition. That would correspond to a very rudimentary level of both individual and collective mathematical comprehension. Poincaré is surely pointing to another type of intuition corresponding to the production of other kinds of images on the basis of which the students (and all of us) think. Let us quote again:

> Sous chaque mot, ils veulent mettre une *image sensible*.… A cette condition seulement ils comprendront et ils retiendront. Ceux là souvent se font illusion à eux-mêmes; ils n'écoutent pas les raisonnements, ils regardent les *figures*. (Poincaré 1999, 105, our emphasis)

I suggest that it is necessary to make here a distinction of at least three levels. In fact, it is not the same (a) to see an apple being cut (sensible intuition), (b) to see the correspondent drawing of an apple cut in several slices (figure), (c) to represent it imagetically, without the presence of the apple or the drawing done with pencil on paper (*imaginative intuition*).

The (a) sensible intuition supposes the presence of a singular, individual entity (be it an apple) or the observation of a determined particular event (the cutting of the apple). The (b) drawing (with pencil on paper) of a figurative representation (a curve, a ring, a circle, a triangle, etc.) corresponds to the second level of that sensible intuition because now what is seen is not the individual in its particularity, but the iconic representation of that individual sketched on paper. There is a huge difference between the individual entity (the apple) and its already more or less symbolic, more or less schematic, more or less abstract representation (the drawing). However, it is yet with the eyes of the body that one sees that figure sketched in the exterior world (the paper). So, we may say that we are still at the level of sensible intuition, even if of diagrammatic nature.

Now, (c) imaginative intuition is the spontaneous capacity for producing an imagetic representation of a circular figure regardless of the presence of any circular object (any apple) and of any drawing (any curve, ring, circular figure). This is a much more complex process since when we see an individual object (sensible intuition) or a specific drawing empirically produced in the external world (sensible intuition of diagrammatic nature), we are always facing to concrete entities, endowed with particular features and dimensions (the drawing of a circle is always a determined entity with determined dimensions, the drawing of a point is always a very small entity with very small dimensions). Now, with imaginative intuition the image is produced by our imagination independently of any empirical entity, just in accordance with the concept of the referred entity. That is what happens in Kant: imagination spontaneously produces the image which will be further synthesized with the concept. The construction of mathematical entities in Kant requires only the pure intuition of time within which imagination produces the image of that entity, that is, that production is made fully *a priori*, without any support coming from experience.

The interesting question now would be: is it possible to make geometry by keeping away from any image, that is, not only without any sensible image (be it an apple or a drawing), but also without any image at all, that is, without any image produced by our imaginative intuition?

For Kant, we cannot. And we cannot because geometry needs images produced by imagination. However, for Kant, what is necessary is neither the sensible image of an apple, nor the empirical, diagrammatic drawing of any round ring, but the image produced by imagination within the pure *a priori* intuition of space and time, more precisely the rule for constructing those images (the *schemata*).

For Brouwer too, we cannot. And we cannot for a much more radical reason: because intuition is the only basis, the very source for the construction of mathematics.

On the contrary, for Bolzano, yes we can make geometry without any images at all. We can and in fact we have to do it because our imagination is unable to produce images corresponding to pure concepts such as mathematic idealities.[5]

Poincaré does not formulate this crucial question, nor does he faces it in its radicalism. We cannot know exactly what his answer would be. At one point, he seems to approach this question when he writes:

> Un grand avantage de la géométrie, c'est précisément que les sens y peuvent venir au secours de l'intelligence…. Malheureusement, nos sens ne peuvent nous mener bien loin. Et ils nous faussent compagnie dès que nous voulons nous envoler en dehors des trois dimensions classiques. (1999, 38-39)

But Poincaré is very ambiguous. On the one hand, he recognizes, against Hilbert, that the sensible intuition may be of the same value in geometry; but on the other hand, he stresses that its role is limited, merely introductory, and only in the context of Euclidean geometry. Even so, let me quote a passage which may give us some insight into his position:

> Nous sommes dans la classe de 4ème; le professeur dicte: le cercle est le lieu des points du plan qui sont á la même distance d'un point intérieur appelé centre. Le bon élève écrit cette phrase sur son cahier; le mauvais élève y dessine des bonhommes; mais ni l'un ni l'autre n'ont compris; alors le professeur prend la craie et trace un cercle sur le tableau. “Ah! pensent les élèves, que ne disait-il tout de suite: un cercle c'est un rond. Nous aurions compris!” (1999, 107)

And Poincaré adds a last brilliant sentence:

> Sans doute, c'est le professeur qui a raison…. Mais il faudrait leur montrer qu'ils ne comprennent pas ce qu'ils croient comprendre. (1999, 107)

The entire quotation is very interesting. But it is to the last sentence that I would like to draw your attention to. What do the students suppose they understand? What do they really understand? What should the teacher make them be aware of?

What I would like to stress is that the answer of Poincaré to the question formulated above—is it possible to make geometry by keeping away from any image, not only any sensible image (an apple or a drawing), but also any image at all—may probably be detected in that last bright sentence.

Students presume to understand the circle in the sensible (diagrammatic) representation they see on the blackboard. But what they do really understand (when they see such figure) is the invisible, ideal circle whose expression may be seen as the inadequate shadow appearing on the blackboard. Of course, the drawing done by the teacher on the blackboard is not a circle. But the drawing done by the teacher on the blackboard is a sensible image of a diagrammatic nature able to “express” the concept of a circle. That is why the students are able "to see the circle" when looking at the drawing done on the blackboard. The students may say (or they may even cry out) that they understand the circle *in* the drawing, but what they really understand is the circle itself (the concept of the circle) expressed *by* the drawing.

Three important consequences may be drawn from the passage given above:

(1) Against Kant, and in accordance with Leibinz, Poincaré safeguards the objectivity of mathematics. (2) Poincare points to another kind of intuition which I propose to designate as *conceptual intuition*, precisely the knowledge of objective mathematical entities. Again, it is Leibniz and not Kant who can be recognized behind this capacity. (3) Poincaré recognizes the expressive capacity of images. Once more, it is Leibniz and not Kant who understood deeply what an image is and the secrecy of its expressive nature.[6]

Unfortunately, Leibniz's awareness of what an image is turns to become impossible after Kant established the rupture between sensibility and understanding. When recognizing the cognitive role of image, Poincaré might suppose that he is retrieving Kant, but it is Leibniz's acute understanding of what an image is, that is at work here. This is what I propose to call Poincaré's conceptual intuition.

Finally, another type of intuition is involved in the process of comprehension of mathematics. It responds to the fact that students (as well as all other men) need to understand the order of the deductive chain:

> Ils veulent savoir, non seulement si tous les syllogismes d'une démonstration sont corrects, mais pourquoi ils s'enchaînent dans tel *ordre*, plutôt que dans tel autre. (Poincaré 1999, 105, our emphases)

It is easy to signalize here another conception of intuition, interior so to say to the demonstration itself, a sudden capacity of seeing at a glance, "*d'un seul coup d'oeil*" (Poincaré 1968, 40), the order of the demonstrative chain.

> Une démonstration mathématique n'est pas une simple juxtaposition de syllogismes, ce sont les syllogismes placés dans une certaine *ordre* et l'ordre dans lequel ces éléments sont placés est beaucoup plus important que ne le sont ces éléments eux-mêmes. Si j'ai le *sentiment, l'intuition pour ainsi dire, de cet ordre,* de façon à apercevoir *d'un coup d'œil l'ensemble* du raisonnement, je ne dois plus craindre d'oublier l'un des éléments; chacun d'eux viendra se placer lui-même dans le cadre qui lui est préparé. (Poincaré 1999, 45-46, our emphases)

We are here facing the well known and much commented *demonstrative or architectonical intuition* (we could say), which enables us to grasp the order of the inferences, the general scheme of the demonstrative movement, or as Poincaré says, "*la marche générale du raisonnement*" (1999, 45).

Poincaré refers to this type of intuition as a sentiment, a delicate feeling hard to define, "*un sentiment délicat, et difficile à definir*" (1999, 46). But what is important to mention is that this demonstrative, architectonical intuition, this intuition of mathematical order—"*intuition de l'ordre mathématique*" (1999, 46)—coincides, at its highest level, with inventive activity. As Poincaré explains, it is because some people are lacking this sentiment that they are unable to comprehend some more elevated mathematics. It is because others have this kind of intuition with some intensity that they are able to comprehend mathematics. And it is because some few others have this capacity in high level that they are able not only to comprehend but to create new mathematics:

> Les uns ne posséderons *ce sentiment délicat, et difficile à définir* … et alors ils seront incapables de comprendre les mathématiques un peu élevées…. D'autres n'auront ce sentiment qu'à une faible degré … ils pourrons comprendre les mathématiques et quelque fois les appliquer…. Les autres enfin posséderont a un *plus ou moins haute degré l'intuition spéciale* dont je viens de parler et alors, nom seulement ils pourront comprendre les mathématiques … mais ils pourront

> devenir créateurs et chercher à inventer avec plus ou moins de succès, suivant que cette intuition est chez eux plus ou moins développée. (1999, 46, our emphases)

Therefore, this demonstrative or architectonical intuition, at its highest level, becomes a special intuition, an "intuition special," that is, an *inventive intuition*, a kind of intuition we will meet again further on.

Let us now come to the *fourth argument* put forward by Poincaré against logicism. We are now facing a psychological argument, mostly formulated against Russell.[7] It concerns the principle of induction which, for Poincaré, is the marvelous result of the activity of human spirit and not, as Russell defended, just the definition of entire number.

Against Russell, who sustained that the richness of mathematics is independent from the (obscure, fragile) power of human spirit, against Russell for whom mathematical induction is not truly an induction because it does not go from the particular to the general since it is always (and already) within the general, Poincaré stresses the powerful virtue of mathematical induction, which, contrary to what happens in other sciences, is able to go from the particular to the general without losing necessity:

> L'induction, appliquée aux sciences physiques, est toujours incertaine, parce qu'elle repose sur la croyance à un ordre général de l'Univers, ordre qui est en dehors de nous. L'induction mathématique, c'est-à-dire la démonstration par récurrence, s'impose au contraire nécessairement, parce qu'elle n'est que l'affirmation d'une propriété de l'esprit lui-même. (1968, 42)

Poincaré also stresses that, unlike the rigorous but sterile syllogistic deduction,[8] mathematical induction allows reaching a conclusion more general than the premises:

> La vérification diffère précisément de la véritable démonstration, parce qu'elle est purement analytique et parce qu'elle est stérile. Elle est stérile parce que la conclusion n'est que la traduction des prémisses dans un autre langage. La démonstration véritable est féconde au contraire parce que la conclusion y est en un sens plus générale que les prémisses. (1968, 34)

Mathematical induction is therefore much more than a series of successive additions as Russell argued. It is the possibility of indefinitely repeating that operation, that is to say, an operative device, a powerful tool of the human spirit endowed with a creative virtue ("*virtue créatrice*," Poincaré 1968, 32), or—as Poincaré explicitly says—an instrument allowing to pass from the finite to the infinite:

> [Le raisonnement par récurrence est] un instrument qui permet de passer du fini à l'infini. (1968, 40)

It is precisely at this point that this (psychological) argument becomes a constructivist *fifth argument*. In fact, for Poincaré, mathematical induction is irreducible, both to the principle of contradiction[9] (on the basis of which we can go on with the analytical development of further and further syllogisms although never creating really new developments) and to the experience, on the basis of which we can go on approaching higher and higher levels of generality but never reach the infinity.

> Ce que l'expérience pourrait nous apprendre, c'est que la règle est vraie pour les dix, pour les cent premiers nombres par exemple, elle ne peut atteindre la suite indéfinie des nombres, mais seulement une portion plus ou moins longue mais *toujours limitée* de cette suite. (Poincaré 1968, 41, our emphases)

But, it is precisely there that, when we recognize the limits of logic (against Louis Coutorat's inflexible logicism) and of experience (against Russell's mathematical empirism), according to Poincaré, we are obliged to recognize that splendid human capacity of enclosing the infinity in one only, single formula. As Poincaré writes:

> Le caractère essentiel du raisonnement par récurrence c'est qu'il contient, condensés pour ainsi dire en une formule unique, une infinité de syllogismes. (1968, 38-39)
>
> Pour y arriver, il faudrait une infinité de syllogismes, il faudrait franchir un abîme que la patiente de l'analyste, réduit aux seules ressources de la logique formelle, ne parviendra á combler. (1968, 40)
>
> Il s'agit d'enfermer une infinité dans une seule formule. (1968, 41)

That is why mathematical induction is considered by Poincaré as an *a priori* synthetic judgment,[10] able consequently to extend our knowledge, an outcome of a very specific type of intuition: a *constructive* (*Kantian*) *intuition*.

That is why, for Poincaré, mathematics is a constructive activity that goes beyond the limits of experience and of analysis. In fact, even if we do not have the capacity of seeing the infinity (with the eyes of the body—by sensible intuition) nor the capacity of imagining it, that is, of spontaneously producing an image of it (by imaginative intuition), however, we have the *a priori* conditions of possibility necessary for its possible iterative construction. As Poincaré writes in a crucial passage:

> Pourquoi donc ce jugement s'impose-t-il à nous avec une irrésistible évidence? C'est qu'il n'est que l'affirmation de la puissance de l'esprit qui se sait capable de concevoir la répétition indéfinie d'un même acte *dès que cet acte est une fois possible*. L'esprit a de cette puissance une *intuition directe* et l'expérience ne peut être pour lui qu'une occasion de s'en servir et par là d'en prendre conscience. (1968, 41, our emphases)

There is here a kind of a process of a *mis en abîme*: the possibility of an iterative construction of infinity is rooted in the intuition of our *a priori* conditions of possibility. The intuition of those *a priori* conditions of possibility offers us the possibility of grasping our power of indefinitely iterating, as if the intuition of our own transcendental structure could open for us (or better, would be the occasion for) the possibility of seeing that infinite possibility. That is to say, against the logicism of Russell and Couturat, Poincaré goes back to Kant.

Poincaré is here in fact very close to Kant[11] and for two different reasons. First, because he stresses not only the capacity of human spirit of being responsible for the construction of mathematics, but the operative nature of that constructability, precisely by mathematical induction (or constructive intuition). As Kant would have said, it is true that we do not have an image of the infinite (our imagination is unable to produce it). But we have the rule for its production (the *schemata*). Similarly, Poincaré's conception of mathematical induction is endowed with an operative, constructive conception of intuition which we proposed to name constructive intuition. Second, because additionally, Poincaré is aware of another kind of intuition—what he calls a *direct intuition*—through which the spirit becomes conscious of its own transcendental capacity for mathematical induction. Let us quote again the last sentence from the passage above:

> L'esprit a de cette puissance une *intuition directe* et l'expérience ne peut être pour lui qu'une occasion de s'en servir et par là d'en prendre conscience [de soi même]. (Poincaré 1968, 41, our emphases)

As Kant would have said, we are here facing the apperception of the pure subject, that is, of the highest and ultimate transcendental structure and of its constitutive role.

Poincaré does not provide a full development of this highest level of direct intuition (or pure intuition, in Kantian terms). But, nevertheless, the awareness of that highest level emerges here and there in his texts. That is what happens, in our view, when Poincaré emphasizes the unity which is necessary for mathematical construction:

> Pour qu'une construction puisse être utile, pour qu'elle ne soit pas une vaine fatigue pour l'esprit, pour qu'elle puisse

servir de marchepied à qui veut s'élever plus haut, il faut d'abord qu'elle possède une sorte *d'unité*, qui permette d'y voir autre chose que la juxtaposition de ses éléments. (1968, 44, our emphases)

As Kant would have said, if the spirit is able to grasp a unique formula for the multiplicity of elements, this is because it uses its own unity for reaching a higher level of construction; because, somehow, our spirit projects its own unity in what it constructs; because its own unity has a constitutive role.

The presence of Kant in Poincaré does not go that far. The awareness of the transcendental root of mathematical unity is a Kantian thesis which, as far as I know, Poincaré never defended. However, it is in line with Poincaré's argument in favor of mathematical unity.

3. Against Formalism

The opponent is now Hilbert, even though Poincaré highly praises the rigorous character of Hilbert's geometry.[12] This does not prevent Poincaré from putting forward two main arguments agains it, each one related to a specific type of intuition.

The *first argument* is developed in a very straight controversial style. Poincaré criticizes Hilbert's program for several reasons: Hilbert does not need to know the meaning of the words and expressions he uses in mathematics; Hilbert does not need to know what the things referred by those words and expressions are; Hilbert only needs a system of linguistic equivalencies; for Hilbert a blind men could be a good geometer, in a word, for Hilbert mathematics might be done by machines. Poincaré does not hesitate to draw a kind of a caricature of Hilbert's program:

> Pensons, dit Hilbert, trois sortes de choses que nous appellerons points, droites et plans, *convenons* qu'une droite sera déterminée par deux points et que, au lieu de dire que cette droite est déterminée par deux points, nous pourrons dire qu'elle passe par deux points ou que ces deux points sont situés sur cette droite. *Que sont ces choses*, non seulement nous n'en savons rien, mais nous ne devons pas chercher savoir. Nous n'en avons pas besoin, et *quelqu'un qui n'aurait jamais vu* ni point, ni droite, ni plan, pourrait faire la géométrie tout aussi bien que nous. (1999, 128, our emphases)
>
> Ainsi, bien entendue, pour démontrer un théorème, il n'est pas nécessaire ni même utile de savoir ce qu'il veut dire. On pourrait remplacer le géomètre par le *piano à raisonner* imaginé par Stanley Jevons…. Pas plus que les *machines*, le mathématicien n'a besoin de *comprendre* ce qu'il fait. (1999, 128, our emphases)

Poincaré's claim is that mathematics cannot be reduced to the manipulation of a sign system. Mathematics is not a merely formal, empty language with its own syntax without any semantic reference to the world.[13]

Now, in this critical claim, intuition is thought by Poincaré as the root for the referential character of mathematics, that is, intuition is thought by Poincaré as a *meaning device*, as the origin for the meaning of words and expressions and for the understanding of what things are (or better, of what their relations are). We are facing here another important issue—the fact that, for Poincaré, the reference of mathematics to the world is not apparent as in Hilbert, but real; even if, for Poincaré, the real means relational (again, the shadow of Leibniz is hovering around).

But now, we have to ask: what might this root be? What valuable advantage may intuition provide for establishing the connection of language to the things of the world, or better to their relations? What might be the role of intuition when faced with a formal system? Why is it impossible to replace the mathematician by the *piano à raisonner* of Stanley Jevons?[14]

In my view, the answer cannot be other than to say that intuition may provide the images needed to connect words and expressions of mathematics to the things and relations they signify. We need images because they allow us to see the things and the relations signified by the words and expressions of mathematics. With Leibniz again, images are cognitive devices.

The *second argument* entails another distinction in Poincare's conception of intuition. As seen above, Poincaré praises Hilbert for the rigorous character of his geometry, namely, he pays tribute to Hilbert's aim of reducing the number of axioms and of making the complete enumeration of the fundamental axioms of geometry.[15] But Poincaré does not believe in the possible accomplishment of this program. He considers it impossible to avoid some infiltration of non-explicit axiomatic suppositions or postulates into mathematical reasoning. And this is impossible because our esprit is active, alive, that is, because intuition has a deep and commonly hidden role in mathematics, precisely at the level of its fundamental axioms.

> Il [Hilbert] voulait réduire au minimum le nombre des axiomes fondamentaux de la géométrie et en faire l'énumération complète; or, dans les *raisonnements vivants*, pour ainsi dire, il est difficile de ne pas introduire un axiome ou un postulat qui passe *inaperçu*. (Poincaré 1999, 129, our emphases)

We are here facing another type of intuition. Let us summarize the situation. Poincaré is not anymore arguing against logicism in its more developed level. Mathematics cannot be taught or learned or constructed on the basis of logic alone, mathematics is a comprehensive task needing sensible intuition in its first steps (apples, drawings), and also spontaneous images produced by a Kantian imaginative intuition. Poincaré is not any more referring to conceptual intuition as objective knowledge of mathematical idealities, nor is he referring to the demonstrative, architechtonical intuition of the order, nor to mathematical induction as a Kantian (constructive) intuition.

Poincaré is now arguing against formalism. But, he is not (as in his first argument) blaming formalism for its lack of intuition as a meaning device. Intuition is not anymore that which avoids the emptiness of words and mathematical expressions. Now, in this second argument against formalism, Poincaré requires a different concept of intuition. Now, he blames the formalist program for its inability to fully explain its fundamental axioms. Now, intuition is thought as a procedure which necessarily infiltrates the foundations of mathematics.

That is, in the first argument against formalism, intuition is a *meaning device*, which gives meaning to words and expressions of mathematics, a root for the understanding of what "things" are, the origin of mental images linking the words and expressions of mathematics to the things and relations they signify, an anti-mechanical remedy which negatively reacts to any attempt to reduce mathematics to a rigorous mechanical procedure.

In the second argument against formalism, intuition becomes a *foundational device*, which infiltrates the foundations of mathematics, an active, energetic procedure, a living force of our mind, able to challenge the full explicitation efforts of a complete formalization.

Of course, it is still necessary to understand what it could be, how it works, this intuition as a foundational, full of life device. Poincaré does not answer this question, even if he frequently emphasizes the vital character of mathematics to which we can only gain access through intuition.

4. Against Conventionalism

There is yet another argument *in the context of Poincaré's peculiar conventionalism*. And this argument (I would like to stress) embraces another important distinction in Poincare's conception of intuition.

Also formulated against formalism, the argument is of a hypothetical form. Poincaré invites us to *pretend* to accept the conventionalist solution according to which the fixation of the main mathematical suppositions (axioms) is done by convention. I quote again Poincaré’s critique of Hilbert’s program:

> Pensons, dit Hilbert, trois sortes de choses que nous appellerons points, droites et plans, *convenons* qu’une droite sera déterminée par deux points et que, au lieu de dire que cette droite est déterminée par deux points, *nous pourrons dire* qu’elle passe par deux points ou que ces deux points sont situés sur cette droite. Que sont ces choses, non seulement nous n’en savons rien, mais nous ne devons pas chercher savoir. Nous n’en avons pas besoin, et quelqu’un qui n’aurait jamais vu ni point, ni droite, ni plan, pourrait faire la géométrie tout aussi bien que nous. (1999, 128, our emphasis)

But—he stresses—even if the conventionalist solution were fully possible, it would be legitimate to go on questioning the origin of such conventions.

> *Admettons* même que l’on ait établie que toutes les théorèmes peuvent se déduire par des procédés purement analytiques, et que ces axiomes ne sont que des conventions. Le philosophe conserverait le droit de rechercher *les origines de ces convenions*, de voir pourquoi elles ont été jugées préférables aux conventions contraires. (Poincaré 1999, 129, our emphases)

Against conventionalists who precisely do not want to answer that question, or better, answer it with arbitrariness, Poincaré wants to find the origin of those conventions. And that origin takes us to an “instinct” which is guiding our choices. As Poincaré states:

> Parmi toutes les constructions que l’on peut combiner avec les matériaux fournis par la logique, il faut faire un *choix*; le vraie géomètre fait ce choix judicieusement parce qu’il est guidé para un *sûr instinct*, ou par quelque *vague conscience* de je ne sais pas quelle *géométrie plus profonde, et plus cachée*, qui seule fait le prix de l’édifice construit. (1999, 129, our emphases)

Intuition is now a “sure instinct” which allows the mathematician to establish the foundations of mathematics not by pure, arbitrary convention, but by choosing those foundations. And the choice is made by an instinct—a guessing as Peirce would say. Intuition is here thought out as an anti-conventionalist device, an instinct enabling the mathematician to choose axioms instead of accepting pure arbitrary convention.

In addition, the choice is made by the mathematician on the basis of a “vague awareness” of a more “deep geometry” capable of giving value to the whole building. How should we understand this deeper level of mathematics in which is rooted the instinct, which allows the mathematician to choose the foundation of mathematics?

Are we here facing an almost *Platonic conception of intuition* as the possibility of accessing the deep, true, eternal roots of mathematics, which would constitute the origin of its foundations? This solution would be in line with Poincaré’s rejection of Stuart Mill’s empiricism, that is to say, in accordance with Poincaré’s defense of mathematical objective (Platonic) idealities. Or, are we facing a kind of *abductive intuition* giving access to a deep and hidden world within which the choice may be reasonably made? Are we facing a possible world in the context of which the choice gains reasonableness, an abductive universe in which, among many others, one specific choice appears as the more reasonable? Peirce would be now the great inspiration.

I suppose that this Peircean solution is more faithful to Poincaré’s thought and more interesting for its operative value. Further, it may better illuminate that foundational, full of life role of intuition which we have mentioned above. In support of this interpretation, it is quite understandable that, precisely in the sequence of the above quotation, Poincaré comes to speak about the instinct as an instrument of invention:

> Chercher l'origine de cet instinct, étudier les lois de cette géométrie profonde qui se sentent et que ne s'énoncent pas, ce serait encore une belle tâche pour les philosophes qui ne veulent pas que la logique soit tout. Mais ce n'est pas à ce point de vue que je veux me placer, ce n'est pas ainsi que je veux poser la question. [Ce que je veux dire c'est que] *Cet instinct dont nous venons de parler est nécessaire à l'inventeur*. (1999, 130, our emphases)

We know that Poincaré explicitly refers to *inventive intuition* in several texts. For instance, in *La valeur de la science*, he states:

> La logique qui peut seule donner la certitude est l'instrument de la démonstration: l'intuition est l'instrument de l'invention. (1970, 37)

Now, that inventive intuition is thought as the "art of choosing among all the possible combinations"[16] and according to several main criteria, experience, pragmatic reasons, architectural and aesthetic principles like simplicity, elegance, symmetry, unity.[17] But this involves—as Peirce would have said—the capacity to choose the best combination.

> Inventer, cela consiste précisément à ne pas construire les combinaisons inutiles et á construire celles qui sont utiles et qui ne sont qu'une infime minorité. (Poincaré 1999, 47)

We are here facing a major issue of Poincare's philosophy of mathematics, namely the role of intuition and the relation between logic and intuition. In the following passage, Poincaré makes this clear:

> La logique nous apprend que sur tel out el chemin nous sommes sûr de ne pas rencontrer d'obstacle. Elle ne nous dit pas quel est celui qui mène au but. Pour cela, il faut *voir le but de loin* et *la faculté qui nous apprend á voir, c'est l'intuition*. Sans elle le géomètre serait comme un écrivant qui serait ferré sur la *grammaire*, mais qui n'aurait pas des *idées*. (1999, 113, our emphases)

This passage is almost the same as in *La valeur de la science* where Poincaré writes:

> L'analyse pure met à notre disposition une foule de procédés dont elle garantie l'infaillibilité…. Mais, de tous ces chemins, quel este celui qui nous mènera le plus promptement au but? Qui nous dira lequel il faut choisir? Il nous faut une faculté qui nous fasse *voir le but de loin* et cette faculté, c'est l'intuition. (1970, 36)

And further on, on the same page, while comparing mathematics with chess game, Poincaré says that what is necessary for understanding the game:

> C'est apercevoir la raison intime qui fait de cette série de coups successifs une sorte de *tout organisé*. (1970, 36, our emphases)

This is a decisive moment. For inventing, we have to choose the best combination (on the basis of Peircean instinctive intuition). But in order to make that choice we have to be able to see the whole—"*voir le but de loin*." And the faculty that teaches us to see is intuition—"*la faculté qui nous apprend á voir, c'est l'intuition*."

For inventing, we need our ability to choose the best combination (*instinctive Peircean intuition*). But for inventing, we also need our capacity for seeing the organic whole ("*le tout organizé*"), that is, we need intuition *as aesthetic visibility*. Without such large visibility of the whole game we would have the rules (the grammar) but not the openness to the real world (the ideas).

> C'est par elle [l'intuition] que le monde mathématique reste en contact avec le monde réel et quand les mathématiques pures pourraient s'en passer, il faudrait toujours y avoir recours pour combler l'abîme qui sépare le symbole de la réalité. (Poincaré 1999, 112)

Poincaré is aware that such intuition (which we have named as aesthetic visibility) can only operate on the basis of a language opened to the world. That is why Poincaré cannot accept Peano's pasigraphy.[18] And that is why Poincaré would have accepted the Leibnizian project of a *Characteristica Universalis*. Indeed, in a very different temporal horizon and much beyond Peano's pasigraphy, Leibinz has looked for the possibility of constructing a philosophical language simultaneously rigorous and meaningful, a universal artificial language enriched by the meaning strategies operating in natural languages.[19] In other words, for Poincaré—as for Leibniz—the aim is to reconcile syntax and semantics, rigor and meaning.

This is obviously a gigantic and impossible project; but yet a desirable one. Let me quote Poincaré one last time, now from *L'avenir des mathématiques*:

> En Mathématiques la rigueur n'est pas tout, mais sans elle il n'y a rien: une démonstration qui n'est pas rigoureuse c'est le néant. Je crois que personne ne contestera cette vérité. Mais si on la prends trop à la lettre, on serait amené a conclure qu'avant 1820, par exemple, il n'y avait pas de mathématiques. (1999, 31)

And, emphasizing again the role of aesthetic intuition, he adds:

> Ce serait manifestement excessif. Les géomètres de ce temps sous-entendait volontiers ce que nos expliquons par des prolixes discours; cela ne veut dire qu'ils ne *voyaient* pas do tout: mais ils passaient là-dessus trop rapidement. Et, *pour bien voir*, il aurait fallu qu'ils prisent *la peine de le dire*. (1999, 31, our emphases)

Hegel describes that "peine" as the patience of the concept. Mathematics needs urgently that patience; to prove what is obvious, to demonstrate what is seen at a glance. What would be mathematics without such intuitive vision? And what would be mathematics without such demonstrative patient procedures?

Philosophy also needs patience. Confronted with the velocity of science, always ready to jump from one truth to another, philosophy is that infinite activity which patiently accepts to return, to revisit, to reexamine, to ruminate, to repeat each day the gesture of Plato's remembrance.

5. Conclusion

To conclude, let me sum up the several concepts of intuition in Poincaré's philosophy of mathematics which identified.

In the controversy with logicism, we have identified (1) a *sensible intuition* as the direct contact with a material entity and with its figurative representative sketched in the exterior world (1a) *diagrammatic intuition*; (2) a (*Kantian*) *imaginative intuition* able to spontaneously produce the images necessary for thinking; (3) a *conceptual intuition* giving us objective knowledge of mathematical idealities; (4) an *architectonical intuition* of the order which, at its highest level, becomes (4a) an *inventive intuition*; (5) a *mathematical induction* as a (Kantian) constructive, operative intuition, an intuition which might additionally be uploaded to a Kantian intuition of the pure subject as a kind of self-awareness Poincaré points to in some passages.

Against formalism, Poincaré thinks intuition both as (6) a Leibnizian *meaning device*, which gives meaning to words and expressions of mathematics, and as (7) a *foundational device*, a living force of the *esprit* (Poincaré is here almost close to Bergson) which makes impossible (and finally irrelevant) any efforts of a complete formalization.

In the context of Poincaré's very peculiar conventionalism, intuition is thought as (8) an *anti-conventionalist device*, an instinct enabling the mathematician to choose axioms instead of accepting pure arbitrary convention. A choice which is made on the basis of the awareness of a deeper level accessible, not so much through a kind of (9) *Platonic intuition* of the true, eternal roots of mathematics, but through (10) a *Peircean abductive intuition*, which, together with our (10a) *aesthetic* (*architectonical*) *capacity of seeing the organic whole*, becomes the root of invention and discovery.

A very last question can now be set up: are all these types of intuition the manifestation of a human constructive capacity, the symptom of a mankind ability for building mathematical objects, the sign of men's power of invention, or are they the expression of the life of mathematics itself? As Poincaré says:

> Il y a une réalité plus subtile, qui fait *la vie des êtres mathématiques*, et qui est autre chose que la logique. (1999, 110, our emphases)

Surely, for Poincaré, it would be impossible to discern, to disclose, to dis-cover, to ex-pose that "life" of mathematics without the help of that human capacity which is intuition. But, what is intuition itself?

Notes

1. In fact, the argument is not very convincing since the example of Cantor could apply for Cantor and not necessarily to all logicism. For a study on the polemics of Poincaré against the logicism, cf. Goldfarb, 1985.

2. As Poincaré writes, "Si les mathématiques n'en avaient pas d'autre [syllogism] elles seraient donc tout de suite arrêtées dans leur dévelopment" (1968, 38). In another passage, Poincaré writes, "Ce n'est certainement pas comme cela que l'esprit humain a procédé pour construire les mathématiques" (1999, 126).

3. "Cette méthode est évidemment contraire à toute saine psychologie" (1999, 126).

4. In fact, Poincaré believes in a non-demonstrated assumption according to which there is an isomorphism between filogenesis and ontogenesis. As he writes, "les questions se poseront successivement à l'enfant, comme elles se sont posées successivement à nos pères" (1999, 112).

5. In fact, for Bolzano imagination does not have the conditions for producing any image corresponding to geometrical concepts and theorems. As he writes, "les lignes que notre imagination parvient à nous représenter, ne sont pas infiniment longues" (2010, 137).

6. For a study on the constitutive role of symbolism in Leibniz, cf. Pombo, 2010.

7. For a study on this controversy, see Heinzmann, 1994.

8. "Le raisonnement syllogistique reste incapable de rien ajouter aux donnes qu'on lui fournit" (Poincaré 1968, 31-32).

9 "… la règle du raisonnement par récurrence est irréductible au principe de contradiction" (Poincaré 1968, 41).

10. "Cette règle, inaccessible à la démonstration analytique et à l'expérience, est le véritable type du jugement synthétique a priori" (Poincaré 1968, 41). This is a major point of controversy with Couturat (1913) for whom there is no place for synthetic judgments in mathematics (cf. Poincaré 1999, 127). For further criticism on Couturat, see also Poincaré, 1999, 153-69.

11. For a close study on the Kantian roots of Poincare's philosophy of mathematics, cf. Folina 1992.

12. For instance, about Hilbert's *Grundlagen der Geometrie*, Poincaré says, "un livre justement admiré et bien des fois couronné…. Voila un livre don't je pense beaucoup de bien, mais que je ne recommanderais pas à un lycéen" (1999, 107).

13. As Poincaré explains, "Croit-on que les mathématiques aient atteint la rigueur absolue sans faire de sacrifice? pas du tout, ce qu'elles ont gagné en rigueur, elles ont perdue en objectivité. C'est en s'éloignant de la réalité qu'elles ont acquis cette pureté parfaite" (1999, 109).

14. Poincaré refers to the Logic Piano constructed by the British economist and logician William Stanley Jevos (1835-82) and exhibited before the Royal Society in 1870.

15. "Ce caractère formel de sa géométrie, je n'en fais pas de reproche à Hilbert. C'était là qu'il devait tendre, étant donné le problème qu'il se posait. Il voulait réduire au minimum le nombre des axiomes de la géométrie et en faire l'énumération complète" (Poincaré 1999, 128-29).

16. "L'art de choisir entre toutes les combinaisons possibles" (Poincaré 1999, 113).

17. Cf., for instance, *L'avenir des mathématiques*, where Poincaré mentions "harmony, symmetry, balancing, order and unity" (1999, 29) and clearly connects these aesthetics criteria with the principle of economy. As he writes, "Cette satisfaction esthétique est par suite liée à l'économie de pensée" (1999, 30).

18. Cf., for example, chapter VII of *Les mathématiques et la logique* (Poincaré 1999, 135-38).

19. For that, Leibniz follows two parallel strategies: (1) logical—to construct a system of signs and operative rules allowing the rigorous expression of thought and its articulations; (2) semantical—to understand the constitutive mechanisms of natural languages which allow their openness to the world. For a developed presentation of the Leibnizian project, see Pombo, 1987.

Works Cited

Bolzano, Bernard. *Premiers écrits: Philosophie, logique, mathématiques*. Ed. Carole Maigne and Jan Sebestik. Paris: Vrin, 2010.

Brouwer, Luitzen Egbertus Jan. "Intuitionism and Formalism." 1912. *Brouwer Collected Works*. Vol. 1. Ed. Arend Heyting. Amsterdam, Oxford, New York: North-Holland, Elsevier, 1975. 123-38.

Couturat, Louis. "Logistique et intuition." *Revue de Métaphysique et de Morale* 21 (1913): 260-68.

Folina, Janet. *Poincaré and the Philosophy of Mathematics.* Hong Kong: Scots Philosophical Club, 1992.

Goldfarb, Warren. "Poincaré against the Logicists." *History and Philosophy of Modern Mathematics*. Vol. X. Ed. Asprey William and Kitcher Philip. Minneapolis: U of Minnesota P, 1985. 61-81.

Heinzmann, Gerhard. "On the Controversy between Poincaré and Russell, about the Status of Complete Induction." *Epistemologia* 17 (1994): 35-52.

Kant, Immanuel. *Kritik der reinen vernunft, akademie textausgabe*. 1787. Berlin: Walter de Gruyter. Trad fr. *Bibliothèque de la pléiade* (Paris, 1980).

Leibniz, Gottfrief Wilhelm. *Die philosophischen Schriften von Gottfried Wilhelm Leibniz*. Ed. Carl Immanuel Gerhardt. Vols. 1-7. Hildesheim: Olms, 1960.

Peirce, Charles Sanders. *Collected Papers*. Vols. 1-2. Ed. Charles Hartshorne and Paul Weiss. Cambridge, Matt.: Harvard UP, 1960.

Poincaré, Henri. *La science et l'hypothèse*. 1903. Paris: Flammarion, 1968.

---. *La valeur de la science*. 1905. Paris: Flammarion, 1970.

---. *Science and Méthode*. 1912. Paris: Éditions Kimé, 1999.

Pombo, Olga. *Leibniz and the Problem of a Universal Language*. Münster: Nodus Publikationen, 1987.

---. "Operativity and Representativity of the Sign in Leibniz." *Studies in Diagrammatology and Diagram Praxis*. Ed. Pombo and Alexander Gerner. London: College Publications, 2010. 1-11.

Philosophy Study, ISSN 2159-5313
June 2012, Vol. 2, No. 6, 398-410

The Stabilizing Role of Material Structure in Scientific Practice*

Anna Estany
Universitat Autònoma Barcelona

Knowledge of the environment is essential for the survival of organisms; but those organisms have to have the capacity to stabilize such knowledge. The aim of this article is to analyze the various strategies for stabilizing human knowledge, with a special focus on its material anchors and their interactions with other stabilization means. In particular, I consider how such stabilization is reflected in scientific activity and practice, and what its repercussions are for the models of science that have dominated the philosophical landscape of the 20th century. My starting hypothesis will be that the role of material anchors in stabilizing conceptual blends is analogous to that of technology in grounding scientific knowledge. The framework I adopt with regard to conceptualization is that of Fauconnier and Turner (2002) on conceptual blends. Just as technology intervenes in scientific practice in conjunction with conceptual elements, so do material anchors, which conjoin other non-material strategies of knowledge stabilization. Endowing knowledge with a material basis may be understood firstly as an element (sometimes a key element) for representing knowledge and offering an explanation, and secondly as a way of providing a scientific hypothesis with empirical grounding. It is this second sense that connects with scientific experimentation and the use of instruments and technology.

Keywords: conceptual blends, material anchors, naturalized philosophy of science, social distributed cognition, conceptual, social and material strategies of stabilization, the role of technology in scientific experimentation

1. Introduction

Knowledge of the environment is essential for the survival of organisms. It enhances adaptability and consequently, improves survival chances. In the case of human beings, knowledge of the characteristics of the environment includes, for example, information on where danger lurks and where to look for food. Such knowledge, however, would be almost useless if it were ephemeral, that is, incapable of being fixed in some way by some means. Organisms therefore have a pressing need for some capacity to stabilize knowledge, store it and, later on, retrieve it.

The aim of this article is to analyze the various strategies for stabilizing human knowledge, with a special focus on its material anchors and their interactions with other stabilization means. From the point of view of the philosophy of science, the question is how this need for stabilization is reflected in scientific activity and practice, and what its repercussions are in the models of science that have dominated the philosophical landscape of the 20th century. Of special relevance for our inquiry is the possibility of establishing a

* **Acknowledgements**: Financial suppot for this research was received from the Spanish Government's DGICYT research project: "Innovation in Scientific Practice: Cognitive Approaches and Their Philosophical Consequences" (FFI 2011-23238). I am grateful to Edwin Hutchins for his support when I was visiting scholar at University of California, San Diego.

Anna Estany, Ph.D., Professor of Philosophy of Science, Department of Philosophy, Universitat Autònoma Barcelona, Spain; main research fields: Philosophy of Design Sciences, Cognitive Approach in Philosophy of Scientific Practices.

relationship between material anchors of knowledge and technology in scientific research. Our starting hypothesis will be that the role of material anchors in stabilizing conceptual blends is analogous to that of technology in grounding scientific knowledge. Just as technology intervenes in scientific practice in conjunction with conceptual elements, so do material anchors, which conjoin other non-material strategies of knowledge stabilization. In both cases the material base constitutes a powerful tool for consolidating knowledge, but the key point here is the connection with the other elements involved in the cognitive processes.

The general framework in which this article falls is a minimalist version of naturalized philosophy of science, which starts from the assumption that cognitive processes occurring in the mind of working scientists, at the laboratory or when writing a paper in front of a computer, are no different from those of other human beings in other circumstances, not even from those of scientists themselves when carrying out everyday private tasks. Very briefly, this is in contrast to the reductionist framework, for instance, Willard Van Orman Quine, or the eliminationist framework, for example, Paul Churchland. It is this assumption of the everyday nature of the cognitive processes involved in science that makes the mechanisms of stabilization and consolidation that are necessary for the conceptualization analyzed in this paper thoroughly relevant to scientific conceptualization. Such mechanisms are a vital part of cognition since the stabilization of concepts involves cognitive processes such as attention and memory.

After this introduction, in the next section I briefly review ideas concerning conceptualization and knowledge. In Section 3, I consider the standard philosophical take on these matters. The following section, Section 4, looks at the shift towards ideas of embodied and embedded cognition over recent years and how the idea of material anchors has come to the fore. In Sections 5 and 6, I examine the so-called method of loci and how space and spatial relations can be seen as a vital ingredient of knowledge. After a discussion of these ideas and some examples in Sections 7 and 8 respectively, I go on in Section 9 to see how stabilization has played an important role in scientific practice and I conclude in Section 10 with some closing remarks.

2. Knowing, Conceptualizing, and Stabilizing

To know implies to conceptualize, and the concepts involved need to be stabilized in order to be useful for us in all the multifarious walks of our lives. That is why conceptualization and stabilization are two key processes in our ability to think. In fact, thinking constitutes a mental activity through which we are able to conceptualize and to know.

In dealing with conceptualization, my framework will be Fauconnier and Turner’s (2002) original idea of conceptual blends inspired by cognitive science. Despite not being the only available approach the authors possess an experimental background in cognitive science which makes it particularly appealing. Furthermore, their model is used by Edwin Hutchins (2005) to introduce the idea of a material anchor of conceptualization. In fact, the approaches of both Fauconnier and Turner, and Hutchins feed into each other. As I point out later, in Section 7, the relevance of a material anchor has to do with the role it plays in stabile conceptualization and this has important consequences for scientific practice.

The theory of conceptual blends deals with concept formation, and is underlain by the idea “that similar general properties of neural binding and simulation lie behind sensorimotor activities, concrete interaction with the world, human-scale everyday experience, abstract reasoning, and scientific or artistic invention” (Fauconnier 2001, 1). According to Fauconnier, conceptual integration (CI) is supported by empirical observation and plays an important role in the construction of meaning in everyday life, the arts, science, and technological development. These possibilities of CI make it thoroughly relevant to scientific research since scientists continually unify concepts and theories.

Here I wish to make a remark on the possible meanings of the term “blend,” since it is not a trivial matter. Most dictionaries—among them the *Oxford English Dictionary* and the *Webster’s Encyclopedic Unabridged Dictionary of the English Language*—define the word “blend” both as mixture and combination. Yet, Gilles Fauconnier’s notion of conceptual blending would be lost if we understood “blend” as mere mixture (i.e., an aggregate of different things in which the components remain individually distinct). The reason is that the result of concept integration is not an aggregate but another concept that possesses an emergent structure that was not present in the concepts it came from. That is the meaning of “chemical blend,” for example, as opposed to that of a mere mixture where substances do not really combine with each other and thus keep their specific properties after the mixture took place.

Conceptual blending operates behind the scenes in the sense that we are not aware of the complexities of perception, for example, when we see a blue cup. Each thought appears to us as if it were simple; yet, from the cognitive point of view, any thought is extremely complex, in the sense that there are several neural processes involved. The model of conceptual blends starts from a structure with some primitives from which we arrive at complex concepts by means of a process of integration of more primitive concepts. Conceptual blending operates in two different mental spaces that work as input spaces to produce a new space: the blend. The partial structure of the input space projects into the blended space, which exhibits the emergent structure. What is of relevance here is that the theory of conceptual blends and mental spaces deals with abstract models of a sort we usually understand as theoretic models. The key idea is the way two or more spaces blend. Different elements from the input spaces are selectively projected into the space resulting from the blending, where new inferences are possible, i.e., the emergent space can be, in turn, the input for a new conceptual blending. Fauconnier and Turner’s (2002)[1] blending strategies result in a blending process that can become ever more complex.

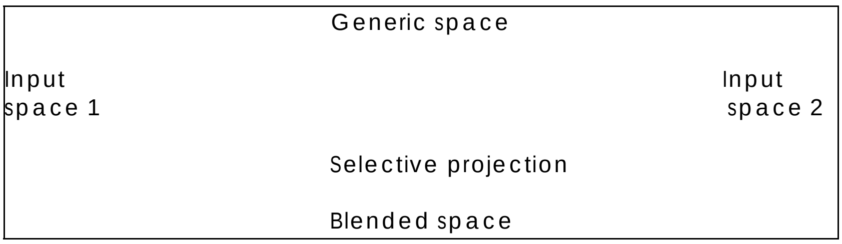

Fig. 1. A Conventional Conceptual Blend.

Source: Adapted from the basic diagram of *The Way We Think: Conceptual Blending and the Mind’s Hidden Complexities* (Fauconnier and Turner 2002, 46).

With regard to stabilization, I will deal here with the stabilization strategies—both material and conceptual—involved in cognitive processes, stabilization strategies that make survival and thought possible. In order to do this, I will make use of Hutchins’s approach (2005) to material anchors on the one hand, and of Robert F. Williams’s (2004) analysis of clocks on the other. In both cases we will be able to discover the main ideas about stabilization strategies, their mutual relations, and the role of material anchors, as a particular means of stabilization plays within this framework. Both Hutchins and Williams admit different strategies of knowledge stabilization; both of them, however, assign a special importance to material ones.

Hutchins notes that there are two strategies for conceptual models to reach stability: one is through what cultural anthropologists call “cultural models”; the other consists of projecting a conceptual blend onto a material structure. Williams, in turn, considers that knowledge stability can be reached by means of conceptual, social, or material strategies.

3. The Philosophical “View”

The relevance of the comments made in the previous paragraphs for scientific practice is obvious, since what scientists have historically done conceptualizes and combines such conceptualizations. What Fauconnier, Turner, and Hutchins offer us is an opportunity to explore the implications of their models for practicing scientists. The scientific equivalents to conceptual blends would be theories or theoretical models, and the blending process would be parallel to that underlying each process of discovery. This view might be questioned from a “discontinuistic” perspective of scientific development, in the sense of non-accumulative, but I think that even in major scientific revolutions, as they are described by Thomas Kuhn in *The Structure of Scientific Revolutions* (1960), new paradigms do not emerge from a theoretical vacuum, but from a set of phenomena that become possible on the basis of new interpretations, techniques, or experiments, as well as from methodological criteria, epistemic values, and even from the contents of other sciences. As a consequence, the general scheme of conceptual blending as a cognitive process adequately fits the way science is done.

Philosophy of science has traditionally been interested in the different ways knowledge is represented and how this reflects and affects our conceptualization of reality; the logical structure of conceptual blends has been of particular importance in that project. As is widely known, philosophy of science focused on the logical structure of the products of scientific research until the decade of the 1960s; and later on the process was thought about essentially from a historical point of view, and not from that of the processes taking place in scientists’ minds. This was what Fauconnier, Turner, and Hutchins studied. Their efforts should be considered as complementary to philosophical analyses of science, not as an alternative to them. Taking this stance does not hinder at all our asking whether studies on the cognitive processes of scientific conceptualization cast doubts upon some of the philosophical models of science or not, a problem whose analysis is beyond the goals of this paper.

4. From Socially Distributed Cognition to Material Anchors

One of the thinkers that have studied material anchors most is Hutchins, whose principal contribution to the field of cognitive science is the idea of socially distributed cognition. This model of cognition restates the unity of cognition, which shifts from the individual to the *interaction* of the individual with some artifacts and other individuals. This idea is close to Hutchins’s proposal on the material anchors of conceptual blends. Even though the latter is a different model from that of distributed cognition and each of them tackles a specific problematic, both models can be seen as complementary, and both assume the relevance of cultural models.

Hutchins takes Fauconnier and Turner’s (2002) idea of conceptual blending as an appropriate framework for characterizing conceptual structure, of which he then seeks the relevant material structure. The association of a conceptual structure with a material structure is a very old phenomenon, and it can be found even among non-human primates. Some important questions on this phenomenon are: (1) How do these conceptual blends get embodied in some material object? (2) Which processes are allowed by such association? (3) How many varieties of blends are possible? (4) What are the cognitive implications of this sort of phenomena? Hutchins tries and explores the possibilities emerging when some or all the structures contributed by one or more input spaces are material. In other words, Hutchins searches for the material anchors either of some of the primitive concepts or of the concept emerging as a result of the blending.

Let us consider a blend where one spatial input is a conceptual domain composed by conceptual items and relations, and the other input is a structured configuration of material elements. Like all blends, conceptual blends have properties that are emergent with respect to those of the input spaces. And perhaps one of the most important of these emergent properties is the capacity to stabilize the representation of conceptual relations. A mental space blends with some material structure that is stable enough to keep the conceptual relation fixed while other mental operations take place. In some cases, the conceptual structures to be represented and handled are so complex that it is not possible at all to endow them with a stable representation by using mental resources only. In these cases, the elements must be fixed (anchored) “somewhere” in order to generate a stable representation of the conceptual elements involved in calculations and enable their manipulation. Such anchorage is achieved by projecting the conceptual elements on some relatively stable material structure. If conceptual elements are projected on a material structure in such a way that the relations between material elements are (wittingly or unwittingly) taken as proxies of relations between conceptual elements, then the material structure is acting as a material anchor. Techniques on conceptual- and material-structure blending alter the amount of cognitive effort needed for the relevant computations. However, a structure is not an anchor by virtue of some intrinsic property it may have. Instead, it is the way it is used that matters. Consequently, material anchoring is context-dependant; then, if material anchoring is functional and contextual, the corresponding cultural legacy will play an important role in the constitution of the material anchor. Roy G. D’Andrade (1989) is one of the cognitive anthropologists who state that cultural models tend to be maintained and strengthened by the behavior and ideas of others. As a consequence, cultural models are one of the elements that contribute to the stability of conceptual blends, but they can also intervene in the other means of stabilization, namely the material basis, which works as an indicator of the interaction among diverse strategies of stabilization.

The cultural process consisting in crystallizing conceptual models in material structures and protecting them from the passage of time puts modern human beings in a world where thinking depends, to a great extent, on the availability of a set of physical structures that can be manipulated (Hutchins 2005). Let us now take a look at two of Hutchins’s examples of conceptual blending with a material base.

5. Method of Loci

This is a very old technique, and can be traced to Aristotle and Cicero. The technique of structuring long stories based on the features of local geography can be found in many non-Western cultures. For instance, in one Baroweni story (Hutchins 1987) a mother travels from the island of Tuma to the village of Osapoula. In the narrative, the mother stops in a series of villages that, when properly located in a map, describe a line from Tuma to Osapoula. Children learn about the geography of the region from stories, and narrators use a version of this method to control the order of events in the narrative. Villages are used as signs to remember in the story.

This method provides a well-known instance of the cognitive use of material structures: in order to remember a long string of ideas, those ideas are put in a given order and associated with a collection of signals from the physical environment in relation to which elements or ideas are to be remembered.

In general terms, input spaces can be said to be comprised by the shape of the movement in a trajectory on the one hand, and by a collection of signals from the environment on the other. Together, these two elements give rise to a blend that consists in the sequence through which the signals "move," and which is an emergent property of the blended space. That space, in turn, may work as an input space for another, more complex blend. In order to remember an ordered string of ideas, those ideas are blended with a collection of signals, generating a space where we can imagine the relevant ideas tied to their corresponding signals. In that space, which is a result of blending, those ideas to be remembered keep the sequential relations already established among the material signals, thus constituting an emergent property of the blended composite space.

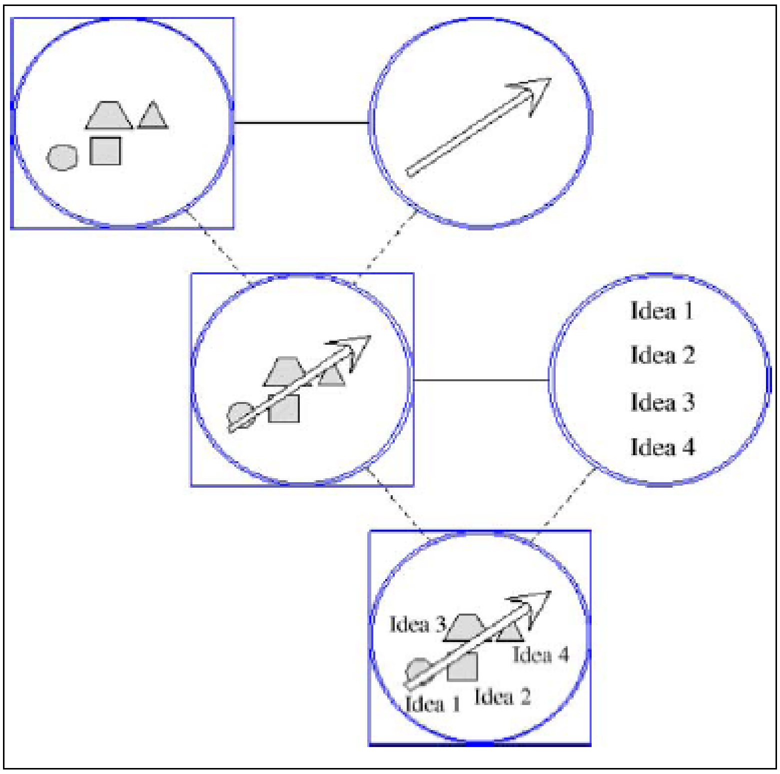

Fig. 2. The Blend Emerged in the Method of Loci (Hutchins 2005, 1564).

For instance, pilots checking their aircrafts before a flight follow a simple trajectory around the aircraft. The shape and direction of the movement impose a set of sequential relations upon the items to be checked. The cognitive problem that is solved by this means is that of controlling action sequentially in order to produce an exhaustive, non-repetitive set of actions.

In considering an example of the use of this technique in scientific practice, different protocols for experiments, both in basic as well as in applied research, can be mentioned. In any laboratory, be it a particle accelerator or a Polymerase Chain Reaction (PCR) in molecular biology, there are protocols designed as rules embodied in material structures that tell us which actions are to be performed.

6. Intelligent Use of Space

The way we organize the objects in our environment is part and parcel of the way we think, plan, and behave (Kirsh 1995). This is the principle guiding all strategies that, one way or another, design the use of space intelligently.

One strategy for dismantling and re-assembling mechanical devices of any kind is to put the pieces drawn from the artifact in a place such that they keep their order they came out of the device, thus representing a sequential order of the objects. When the device is re-assembled, we follow the same order, but in the opposite direction.

Another example of intelligent use of space is a queue. This cultural practice creates a spatial memory of the order of arrival of customers. Subjects in the queue use their own bodies and their relative position in space to codify order relationships. The practice of making queues fits in the ecology of other practices that are part of the cultural model "first come first served." For instance, consider queues of passengers at check-in desks in an airport. Only one line of clients waiting to be served at several check-in desks constitutes a spatial structure that fits the mental-cultural idea "first come first served" better than the structure of one line at each check-in desk.

Similarly, space distribution in laboratories is significant to research success. The environment can favor (or hinder) communication among the members of the research team, as well as the relationships among director, researchers, graduate students, and technicians. In sum, the structure of a rather hierarchical organization (or network) will be reflected in spatial distribution. But we may construe that relation the other way round: a given spatial distribution can favor increased collaboration or increased hierarchical verticality among the members of a team.

7. Discussion

Taking into account the goals of this paper, Hutchins's ideas emphasize material anchors, but they can be detached neither from the framework of socially distributed cognition, nor from the importance of cultural models in cognition. Thus, a material means becomes an anchor for a conceptual blend.

The central place given to material structure should also be taken as a way to compensate for the scant importance it had in the symbolic paradigm on information processing, which dominated the earlier stages of cognitive science. It seems to me that we should not think that material anchors are the only way of knowledge stabilization, or that they are not related to any other way of stabilization. It is also important to note that anchors (even material ones) admit degrees of strength; they can be stronger or weaker depending on technological and cultural factors. As will be shown below, this point is essential for scientific practice, because it affects one of the strategies of representing knowledge, namely language (natural or mathematical). In summary, material anchors are a major source of knowledge stability; this idea is supported by empirical results in cognitive psychology and has significant implications for scientific practice, since one of the aims of scientists is to consolidate—not only to create—knowledge.

8. Clocks: An Example of Convergence of Material and Cultural Anchors

In his doctoral dissertation ("Making Meaning from a Clock: Material Artifacts and Conceptual Blending in Time-Telling Instruction"), Williams (2004) applies Fauconnier and Turner's and Hutchins's models to time-telling learning. Williams aims at exploring phenomena that go beyond the internal/external division usual in cognitive models, thus providing a more complete and consistent approach to human cognition. Williams belongs in a new paradigm of cognitive science, which views cognition as embodied, embedded, situated, and dynamic, all of which imply a new cognitive semantics where the construction of meaning depends on the embodied experience, the mental simulation, and the extension to abstract domains through analogy, conceptual metaphor, and conceptual blend (Williams 2004, 2).

Even though he admits the importance of material stabilization, Williams advocates other strategies for stabilization, such as conceptual and social ones. According to Williams, these three strategies of stabilization cannot be viewed as independent from each other, but should be seen as being related and engaged in constant interaction.

Among the conceptual sources of stability, Williams mentions the ways humans structure information. Work on significant chunks (Miller 1956), scripts (Shank and Abelson 1977), conceptual frameworks (Minsky's frames 1995), "Idealized Cognitive Models" (ICM, Lakoff 1987), and "cultural models" (D'Andrade 1989) should be included here. All these are means for structuring knowledge and keeping it stabilized after the instant it emerges.

One argument supporting these ways of structuring information is based on the fact that human beings share the same bodily configuration and inhabit the same world. This fact provides support for the assumption of the existence of universal mechanisms of knowledge stabilization, provided circumstances are "normal." However, conceptual models are influenced by the activities of the cultural group and by the meanings its members create within the group. In other words, the stability of knowledge which is achieved through models and frameworks due to individuals belonging to the same species has a universal basis; yet, there is also variation due to specific individual and social characteristics. And here social strategies of stabilization, among them cultural models linked to cultural practices, come into scene.

The importance of clocks consists in that telling time implies that a person interacts with an artifact (a clock), and this provides us with an opportunity to explore the relations between conceptual models and material artifacts in a cognitive activity. We should not forget that the interaction between individual and artifact is a fundamental feature of the cognition unit in Hutchins's model. Besides, learning to tell the time requires instruction, a fact that provides an opportunity to observe, register, and analyze social interrelations (another element of Hutchins's unit of cognition).

This case study attempts to show how conceptual and material structures interact in the cognitive process of telling the time; where practices, artifacts, and conceptual models come from; and why they are as they are. This will give us an idea of how we can tackle errors in time-telling, and then, what are the differences in conceptual understanding between novices and experts.

We will not examine the details of Williams's research, which is rigorous and wide-ranging, in addition to being supported by some of the proponents of the cognitive models it applies.[2] However, it is relevant for our inquiry to make reference to the conclusions reached by Williams with regard to cognition in general, and above all, with regard to stabilization strategies.

Williams considers that cognitive functional systems come from cultural models and their anchorage in artifacts, practices, etc.. This implies that instruction is focused in perpetuating functional systems that are positively valued by the cultural group. Such oriented conceptualization amounts to using the semiotic resources available—speech, body, and world—in order to give novices the power to build relevant meanings in the course of the activity.

In this framework, language plays an important role, but it is only one of the several strategies available for building meanings, since gestures and artifacts build meaning, too. What counts as information may emerge only from the interrelation among gesture, environment, and speech, rather than from what each of them represents on its own.

If there is scant or nil material support, the cognitive system can replace it with familiar structures filled out with imagination or through mechanisms of pattern recognition. In the clock example, the case would be that numbers in the sphere are not marked (this is quite usual nowadays, due to design and fashion motives). Novices have less experience and have their conceptual models of clocks and time-telling less developed than experts. These circumstances make it more difficult for the former to fill the gaps and maintain conceptual relations without the stability provided by the material basis (i.e., the marked numbers).

There are important differences between stabilizing knowledge through some material structure and doing so through conceptual models. In the first case we have a physical entity (fixed, even though it may have mobile parts), and perceptual space is primarily oriented by sensorial data coming from the physical world. In the second case, we do not have such physical structure, and perceptual space is primarily oriented by activity in the sensory-motor areas of the brain. We may conclude that we have more means than those that are only material for stabilizing knowledge, but there is no doubt that, in general, it is material ones that provide the most robust stabilization. We should not forget that mental images are always less vivid and persistent than sensory perception, though mental images of very familiar objects and environments may possess enough detail and immediacy in order to work as a stabilizing basis for many cognitive activities.

From the clocks example, we can draw some conclusions that are significant with regard to scientific practice. In the first place, the example shows the convergence of different stabilization strategies. Second, it is a cognitive model of learning, as we find in the relationship between senior researchers and graduate students. Finally, learning time-telling proceeds by means of an artifact (the clock) in the same way that graduate students learn to handle laboratory instruments and techniques.

9. Conceptual Stabilization in Scientific Practice

Scientific research consists essentially in knowing and explaining the world, conceptualizing it, and stabilizing those conceptualizations. Consequently, all that has been said in the preceding paragraphs is of importance to scientific practice. In fact, every scientific theory is a conceptual model—though the converse is not true—at least if we abide by the more or less technical characterizations of “theory” provided by philosophers of science. At this point, it is interesting to see how we relate ideas on conceptual models coming from cognitive science with available ideas on scientific theories.

From Hutchins’s point of view (and we may say, from the cognitive point of view), there are constraints among the elements that compose conceptual models, i.e., requirements as regards the relations between conceptual elements. In order for them to play an important role in the reasoning process, conceptual models must be cognitively stable. And in order to achieve stability, a conceptual model has to maintain its system of constraints at the same time the model is subject to mental (or physical) manipulation. Stability may come from different circumstances. In some conceptual models, stability comes from simplicity, in others that are more complex, stability comes from their being embedded in conventional (culturally shared) internal mental structures that are well learned and automatically applied. A conceptual model with those properties is a cultural model. With sufficient practice, support and motivation, even complex, rather arbitrary models can be learned. In a system of cultural models, each model constrains and is constrained by a wide network of relations with other models. It is possible to reason efficiently on complex problems when they are expressed in terms of cultural models. Other conceptual models achieve stability through being embedded in an external physical object, for example, pilots checking their aircraft or the use of space in a queue, already pointed out. Those problems too complex to be kept in memory as cultural models or internal conceptual models can be expressed and manipulated in an external object or structure (Hutchins 2005, 1556).

In the case of scientific practice, the acquisition of conceptual/cultural models would correspond to what we know as scientific instruction, which could be rightly called "scientific socialization"; hence the importance, already noted by Kuhn, of text books. Now, talking about cultural models of scientists requires a clarification. In general, this talk refers to personal and social differences, with indicators such as sex, religion, ethnic origin, etc., that according to constructivists determine the sort of science that is made, and not only this but also the acceptance and epistemological value of theories. However, even though a given cultural baggage exerts an influence on scientific practice, we must remember that for a person to become a researcher, she needs to pass a series of instructional stages where, even when they do not make "intellectual clones," they do homogenize and polish personal and cultural characteristics. A different matter altogether is whether this is an advantage or a disadvantage, or whether it hinders or benefits scientific research; in any case it is a fact we should not forget.

All this makes me think that when speaking about cultural models in scientific practice we cannot formulate them in terms of indicators such as those previously mentioned (sex, religion, ethnic origin, etc.) only. By this I am not saying that cultural models in the sense construed by anthropologists do not influence scientific practice and manifest through biases, both when choosing the research matter (the less distorting effect) and when reasoning about it.[3] But scientific culture *strictu sensu* should be understood as "the rules of the game" on the one hand, and as theoretical models for undertaking a given inquiry on the other.

As to the rules of the game, the main cultural model is the methodological model, whose boundaries may be very imprecise at the beginning, but which becomes progressively more concrete—even rigid, to some extent—in the form of very precise rules and procedures. Apart from methodology, the ontology and metaphysical principles underlying any scientific practice are models of scientific culture equivalent to the cultural models studied by Hutchins and D'Andrade, and are of great help for our everyday activities, as well as a source of cognitive effort saving. Also, theoretical models or theories are a part of cultural models that have stabilized through scientific instruction or socialization.

At the beginning of this paper I noted that the main strategy for providing a material basis to knowledge is through technology (both techniques and physical instruments). Here I return to the idea of material anchors as a means for stabilizing ideas. If we consider the role of technology in scientific research, it not only plays a stabilizing role, but may also act as a support for explanation, learning, and the empirical grounding of knowledge. This is particularly important for a philosophy of experiment (also referred to as experimental traditions), since experimentation is considered at least at the same level of theorization.[4]

The case of learning how to "read" Magnetic Resonance Imaging (MRI) studied by Alac and Hutchins (2004) provides a good example of explanation and learning from conceptual models and material anchors. The idea is that when novices learn how to read MRI, not only words and explanations from experts have a role, but also graphics and gestures do, and that it is the collection of all these elements that will allow the effective reading of the images.

The history of psychology too offers an example of how material anchors provide empirical grounds for a series of theoretical assumptions. Let us go back, just for a moment, to the first half of the 20th century, when the behaviorist paradigm was still dominant in the field of psychology. Behaviorists abandoned the study of mental phenomena because these could not be observed. With the cognitive revolution, however, where one of the disciplines in the theoretical framework was information science, certain mental phenomena could be simulated on the computer. Even though it was an analogy or a metaphor, the study of how humans processed information now had a material basis that gave the endeavor credit and scientific grounding. Some thinkers, such as Arthur L. Blumenthal (1975; 1979), said that cognitive psychology was a return to Wilhelm Wundt and his frequently overlooked founding view that psychology is the science whose objective is the study of the elements of consciousness. It may well be in some aspects but it is not emphatically the case in some others. That is to say, cognitive psychology put back on the scene phenomena that had been banned by behaviorism, but the methods used and the material resources available for representing knowledge were not the same in one and the other. Let us see, then, a case where a given technology, computers, allowed knowledge to be consolidated that would otherwise have remained in the realm of hypotheses if not of mere speculation.

Finally, when dealing with stabilization strategies and material anchors, we cannot dispense with saying something about language—or, better, languages—and their stabilizing capacity. If we take into account that an important part of scientific knowledge is expressed through natural languages (even mathematical models have replicas in linguistic propositions), we can understand the importance of assessing the degree of stability that language provides to knowledge and see whether it constitutes a genuine material anchor. Fauconnier and Turner think it does, but the issue requires some reflection on our part, and cannot be solved by means of a Solomonic middle course. One word *can* be seen as a material anchor for a conceptual blend, but then contribution of the material means to the blended space would be minimal, though linguistic units go beyond vocabulary, they may imply more interesting material structures.

Hutchins too wonders to what extent language (spoken and written) can be taken as a material anchor, and notes that "[m]aterial structure can have many kinds of relations with conceptual structure. One of the most common is the relationship of representation. All of language depends on this sort of relationship so this one comes easily to mind for linguists" (2005, 1556). Yet, he notes firstly that there are more basic principles by means of which meaning is attributed or projected into a material structure, and secondly, that although language may be considered a weak material anchor, particularly lexicon, linguistic units larger than words can be more important material structures.

Having said this about language, when it comes to scientific practice, the problem of the meaning of mathematical language, from the point of view of stabilization, arises. Can we apply Fauconnier's and Hutchins's opinions on the stabilizing capacity of language to mathematical language? Galileo Galilei says in *Il Saggiatore* (1623) that nature is a book written in the language of mathematics, but is such writing a material anchor? Only if we think in a scientific culture may we say that mathematical models can be material anchors to mathematics, at least to the same degree that natural language can be a material anchor for any human being.

All this makes us rethink the relationships existing among different stabilization strategies. It could be that characterization of one strategy of stabilization or another is a matter of degree and center of gravity. However, in any kind of stabilization, several sources always converge, though they intervene in different degrees. The counterpart of such circumstance in scientific practice would be that intervention and the weight of theory and experiment are also a matter of degree and emphasis, and that there is no reason to consider it as uniform or taking place in every discipline, every historic period, nor in every stage of a research.

10. Closing Remarks

Stabilization is essential for supporting one of the pillars of scientific research, namely, that knowledge is cumulative, though epistemic progress is not necessarily linear. This may seem trivial and to be taken for granted, but here we have dealt with the meaning of stabilization as a cognitive process.

Knowledge stabilization can be achieved through several strategies that can be categorized as conceptual, cultural, and material. Without underestimating the importance of the first two, we can say that material strategies of stabilization constitute very powerful anchors that humans have historically used in the course of evolution. However, this taxonomy of sources is blurred in reality, both in everyday actions and in scientific practice. Consequently, what we really find is the interaction among different strategies of stabilization.

Endowing knowledge with a material basis may be understood firstly as an element (sometimes a key element) for representing knowledge and offering an explanation, and secondly as a way of providing a scientific hypothesis with empirical grounding. It is this second sense that connects with scientific experimentation and the use of instruments and technology.

The case of the interpretation of MRI and that of the history of psychology illustrate that the material basis consists not only in giving shape to knowledge but in representing and transmitting it. The set that makes interpretation possible is constituted by the graphics, gestures, and theoretical models from neurobiology.

We must not forget the importance of context, since material anchors are not so by virtue of some intrinsic property, but of what they represent and of the meaning they have in a given context. This fact links them to cultural models, though later on such models are expressed in some material way.

Notes

1. See chapter 7 of the book, where the authors put forward a classification of networks as a function of their conceptual complexity. From less complex to more complex, they are: simplex, mirror, single-scope, and double scope.

2. The tribunal evaluating Williams's dissertation was composed of Gilles Fauconnier, Edwin Hutchins, Charles Goodwin, Roland Langacker, and Rafael Núñez.

3. David Kitcher (1993) has analyzed these biases.

4. We should include the approach put forth by philosophers of science such as Ian Hacking, Peter Galison, Josef Rheinberger, Sergio Martínez, and Javier Ordóñez, among many others.

Works Cited

Alac, Morena, and Edwin Hutchins. "I See What You are Saying: Action as Cognition in fMRI Brain Mapping Practice." *Journal of Cognition and Culture* 4.3 (2004): 1-30.

Bache, Carl. "Constraining Conceptual Integration Theory: Levels of Blending and Disintegration." *Journal of Pragmatics* 37.10 (2005): 1615-35.

Blumenthal, Arthur L. "A Reappraisal of Wilhelm Wundt." *American Psychology* 30.11 (1975): 1081-88.

---. "Wilhelm Wundt, the Founding Father We Never Knew." *Contemporary Psychology* 24.7 (1979): 547-50.

D'Andrade, Roy G. "Cultural Cognition." *Foundations of Cognitive Science*. Ed. Michael I. Posner. Cambridge, Mass.: MIT, 1989. 795-830.

Fauconnier, Gilles. "Conceptual Integration. Emergence and Development of Embodied Cognition." *Proceedings of the Emergence and Development of Embodied Cognition Workshop. Beijing, China, August 2001*. 22 Feb. 2007. <http://www.ifi.unizh.ch/ailab/people/lunga/Conferences/EDEC2/invited/FauconnierGilles.pdf/>.

---. "Fusión Conceptual y Analogía." *Cuadernos de Información y Comunicación* 10 (2005): 151-82.

---. *Mappings in Thought and Language*. Cambridge, UK: Cambridge UP, 1997.

Fauconnier, Gilles, and Mark Turner. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books, 2002.

Hutchins, Edwin. "Material Anchors for Conceptual Blends." *Journal of Pragmatics* 37.10 (2005): 1555-77.

---. "Myth and Experience in the Trobriand Islands." *Cultural Models in Language and Thought*. Ed. Dorothy Holland and Naomi Quinn. Cambridge, UK: Cambridge UP, 1987. 269-89.

Kirsh, David. "The Intelligent Use of Space." *Artificial Intelligence* 73.1-2 (1995): 31-68.

Lakoff, George. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: U of Chicago P, 1987.

Miller, George A. "The Magical Number Seven, plus or minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63.2 (1956): 81-97.

Minsky, Marvin. "A Framework for Representing Knowledge." *Computation and Intelligence: Collected Readings*. Ed. George F. Luger. Menlo Park, CA: AAAI, 1995. 163-89.

Norman, Donald A. *El aprendizaje de la memoria*. Ed. Alianza Madrid. Trans. Sebastián Gascón, María Vicenta and Tomas del Amo. N.p: n.p., 1985.

Schank, Roger C., and Robert P. Abelson. *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

Williams, Robert F. "Making Meaning from a Clock: Material Artifacts and Conceptual Blending in Time-Telling Instruction." Diss. University of California, 2004.

Philosophy Study, ISSN 2159-5313
June 2012, Vol. 2, No. 6, 411-416

Wittgenstein and Kuhn on Paradigm

Ines Lacerda Araujo
Federal University of Parana

Wittgenstein in *Philosophical Investigations* and *On Certainty* develops a revolutionary conception of knowledge, of philosophy of language, and of philosophy of science. Very close to language-games is a concept that is important but did not attract much attention, the one of paradigm. In this paper this concept is analyzed together with notions such as context, use, reference, language-games, reality, science, truth, method, propositions, among others by means of which a new view of epistemology emerges. As for Kuhn, paradigm is his most known notion and his biggest contribution to innovate the history of science. I contrast these two conceptions and conclude that they have different uses, and that some difficulties that emerge from Kuhn's concept of paradigm would be dissolved if we take into consideration Wittgenstein's approaches of language and the uses it has in our life forms.

Keywords: epistemology, Wittgenstein, Kuhn, paradigm, philosophy of science

1. Introduction

The concept of paradigm is essential to the philosophy of science. The first philosopher to use this concept in its epistemological sense was Ludwig Wittgenstein in his late writings, when the focus changed from formalism to the ordinary language, the language-games, and epistemology. This paper will argue that there are hints of an epistemology centered on concepts like paradigm, context of use, science as a practice, the role of empiric propositions and of the hinge propositions. As for Thomas Kuhn, the concept of paradigm is central to his conception of the history of science with a root in Wittgenstein's language-games.

2. Wittgenstein's Epistemology and the Concept of Paradigm

Empiric propositions have functions in our forms of life. Some of them are used to represent facts as they are disposed by means of scientific knowledge; others are used to make assertions on how things are. They have specific uses, they cannot be considered anymore as figuring states of affairs by means of logic symbols. In both *Philosophical Investigations* and *On Certainty*, it is shown that logic is part of ordinary life practices, and meaning something to someone is a process in which enters reference that is not isolated from context, from the normal situations of use of the language-games. Words are used as we use a tool and the meaning does not depend on naming objects.

Language is a kind of action and behavior. It permits us to invent stories, to measure things, to count them, to make an announcement, to judge, to teach, to draw a sketch, to learn directions on a map and many other activities that the grammar makes possible. Grammar to Wittgenstein is to be understood as any rule that guides or creates the possibility of expressing concepts and thoughts, of formulating judgments on things and on the

Ines Lacerda Araujo, Ph.D. in Linguistic Studies, Emeritus Professor of Philosophy of Language, Federal University of Parana, Brazil; main research fields: Philosophy of Language, Epistemology.

immense variety of situations. There is no hub and no essence of language or thought. To name is not a mysterious mental process, it is a mean of representation. Grammar can provide *paradigms* (*Vorbilden*) that are instruments of the language to make comparisons, to teach, to give things a meaning, to communicate desires, intentions, messages, etc..

There are different criteria to know if and when a proposition refers to, what observations are taken into consideration when it is used, which are not. In other words, it is simply necessary language, rules of a grammar, certain paradigms of grammar such as long or short sentences to give an order, for instance, “Bring me a slab” or just “Slab” (Wittgenstein 2001, §20). Our grammar has different paradigms. The difficulty to mean something or to name something that does not exist (as an object or as individual or even as a concept in mind) disappears. Does the standard meter-rule in Paris really exist? No, it has a “peculiar role in the language-game of measuring with the meter-rule” (Wittgenstein 2001, §50). If this thing did not exist, it would not have a name, means simply this, “If this thing did not exist, we could not use it in our language-game. What looks as if it *had* to exist, is part of the language. It is a paradigm in our language-game; something with which comparison is made … but it is none less an observation concerning our language-game—our method of representation” (Wittgenstein 2001, §50).

The correspondence between an element of language and things follows paradigms, just as a technique to correspond colors to their respective squares, each one receiving a name, which shows the necessity of a convention. There are other ways to proceed with colors, for instance, to distinguish between them, to choose a darker blue, or a reddish yellow, to put into words the green of a landscape, and so on. “An example of something corresponding to the name, and without which it would have no meaning, is a paradigm that is used in connection with the name in the language-game” (Wittgenstein 2001, §55). It is possible that one looses the skill to make comparison between colors, and this means that a paradigm that was used in his language is lost. Paradigms are models used in different behaviors and activities that are accepted and practiced, that govern our behaviors and intentions, and that guide the use of concepts. The concept of “pain,” for instance, is learned when a language is learned. And the use of concepts is learned under certain criteria that are guided by a paradigm. Whenever a paradigm varies the use of the concept, the criterion also varies.

Wittgenstein gave up the project of linking language and the world, of attaining the essence of them both. There is no unique model for language to assert how things are and there is no common foundation to each and every language-game. Logic belongs to the ordinary language, so to describe a language-game is something that belongs to logic. To understand a sentence is a matter of rules, uses, habit, techniques that are learned and that have been developed in mankind cultural evolution with the changing of the institutions. The relations between human needs and nature create habits, beliefs, and behaviors according to the geographic, physic aspects of the world and the environment. There is no mental or private reign above normal circumstances. Our “mental” processes such as attention, understanding, imagination, memory are related to human needs and have specific uses. They respond to objectives that produce different kinds of learning. We have learned to calculate, to measure things, to draw projects, to make stories, to read, to write, and all these varied activities depend upon language, grammar, and paradigms.

These philosophical conceptions of language gave conditions to a *new kind of epistemology*. Wittgenstein’s conception does not fit into traditional conceptions of knowledge as a unity or as having one solid and unique foundation that each conception has to take into account, such as ideas to idealism, experience to empiricism, the existence of things independent of the subject to realism, the impossibility of truth to skepticism. Wittgenstein thinks that there are bases, but they are simply related to material and cultural features, to anthropologic aspects, to psychological conditions, to capacities such as learning language. And to learn a language is not a mysterious process, it is similar to the processes of learning behaviors. To mean, to see, to know, for instance, depend on the context and on the circumstances of use. To see an aspect is different from seeing something as this or that. All these are grammars.

Concepts belong to a system of presuppositions. To verify an assertion is not the general criterion for science or the paradigm of knowledge. Verification is one process among many others to obtain knowledge, it is necessary that every process be justified by need and uses. The truth of empiric propositions is based on a reference frame and not on immediate visual experiences. Without context and without taking into account normal circumstances, the propositions have no meaning.

3. Certainty and Scientific Knowledge

There are propositions that need not to be verified. Wittgenstein called them “hinge” propositions. This does not imply that they are on the foundation of all kinds of knowledge. There are some things that we know, when one says “I know this is a tree,” the situation may be two people discussing on the differences between bushes and adult trees. A question or an assertion must make sense, such as someone saying that he or she has reasonable bases to make an assertion. And this implies another person that is capable of knowing what has been taken into consideration in this case and that someone knows such or such thing. This conception eliminates implications of a real object that has to be present to both speaker and hearer. It is not the presence of an object that is considered the source of certainty. Certainty belongs to different situations and conditions. There are things about which there is no sense to doubt, for example, the very ancient age of the earth. Doubt in these contexts would modify the order of things. What memory, history, investigations have proved up to now have been successful, useful, and worthy to be given.

There are convictions accepted because they have practical effects that anyone can use and rely on. They need not to be investigated, they are accepted whenever there are changes that require revision on the way things are disposed in the new situation. How a fact enters someone’s conscience? What occurs in the external world, it is said that this is not known unless it is considered as sense data. Wittgenstein disagrees:

> This would give us a picture of knowing as the perception of an outer event through visual rays which project it as it is into the eye and the consciousness. Only then the question at once arises whether one can be *certain* of this projection. And this picture does indeed show how our *imagination* presents knowledge, but not what lies at the bottom of this presentation. (1991, §90)

The problem is that one cannot be sure if this image corresponds to the exterior sense data, so this is an issue of how one imagines knowledge, and this is not what is at the “bottom” of this way of presentation. This way of presentation requires a connection between mind and world that is the source of all kinds of knowledge. But Wittgenstein thinks that there is not only one process of knowing by means of a necessary link subject-object. To verify, to assert things, to name, to describe, to be sure of, all these are different ways that are used in our life forms. There is a reference frame that can be used in different cultures to make inquiries: how they are transmitted, what is the role of scientific research, how a system of propositions is required to verify statements. And all this varies.

When someone states that he knows, this has to do with strong reasons that are learned together with worldviews whose truth may or may not be demonstrated. In order to make objective investigation empiric, propositions are used and these propositions that have been learned constitute nets of an inquiry system. “We are quite sure of it does not mean just that every single person is certain of it, but that we belong to a community which is bound together by science and education” (1991, §298).

Scientific knowledge forms a system of convictions, it is supported by experience, but this is not viewed as a foundation or prerequisite, it is simply the way science functions and its knowledge is transmitted; people believe in facts of geography, chemistry, physics, and their convictions and beliefs rely on these kinds of knowledge. But facts do not speak for themselves. It is necessarily an adjustment with a whole system of propositions and with actions. The changes of situations have a crucial role in the use of language-games. What is the appropriate language-game, the context, the use, the situation will decide. Knowledge is not a mental process, but a way of acting. These conceptions are very close to pragmatism. Indeed, as Wittgenstein considers language as a kind of behavior or a way of acting, and as he takes away reference to things of the external world as the main function of language, his approach is comparable to some of Dewey’s ideas on knowledge. To both Wittgenstein and Dewey knowledge is part of the ordinary tasks and activities, it is not a mental capacity and so the mind is not the theatre where ideas are projected.

Wittgenstein approximates science of life and human needs. In *On Certainty* he says about certainty on the science domain: “That is to say, it belongs to the logic of our scientific investigations that certain things are *indeed* not doubted” (§342).

So, knowledge has different and diverse ways of usage. It has solid points of departure, but this does not mean that there is only one criterion or one and only specific mode of proving, or that certainty belongs to the territory of science, of proved facts. What we know is a matter of language-games and not a privileged relation between subject and object, between a mental representation and a represented world. This is a picture that corresponds to traditional conceptions such as idealism, realism, and empiricism that presuppose that mind and world are somehow related. They end up in enigma, philosophical problems that have to be dissolved.

So this is a revolutionary and critic epistemology. Kuhn was interested in this new way to conceive knowledge—speciality, language-games, paradigm, and the notion of incommensurability of scientific theories. The concept of paradigm will have a crucial and main role although more specific. It will solve the problem of the history of science, as what I will analyze below.

4. The Concept of Paradigm of Kuhn

Richard Rorty wrote that Kuhn’s “*The Structure of Scientific Revolutions* owed something to Wittgenstein’s criticism of standard epistemology” (1979, 322) and proceeded in a “fresh way.” Newtonian physics is not paradigm for knowledge as Kant advocated. There is no algorithm nor neutral scheme that philosophy of science would have to construct in order to choose among scientific theories, there is not a common ground where culture could find definitive, cognitive, or rational criteria for knowledge, argues Rorty.

Kuhn emphasizes that paradigms are similar to Wittgenstein’s concept of “family resemblance,” that is, a language-game does not attribute a name with basis on a closed “set of characteristics” that the object would have and that all needed to be included by the name. “Closed historical investigation of a given specialty at a given time discloses a set of recurrent and quasi-standard illustrations of various theories in their conceptual, observational, and instrumental applications. These are the community's paradigms, revealed in text books, lectures, and laboratory exercises" (Kuhn 1996, 43).

This conception of science is historicist and instrumentalist, it is based on context, not on some absolute view of science. There is progress because there are paradigms, that is, assumptions, beliefs, and techniques to select, evaluate, and criticize a whole field of research, and this is not implicit and ready in facts. Facts are selected by scientists that share a paradigm in normal science. Normal science functions as a kind of puzzle solving problems. Without a paradigm there is no science, the field of research is configured by the paradigm that is a pattern to guide science toward progressive knowledge and the production of instruments, and it also permits prevision. A paradigm does not solve every problem, there are gaps and questions that cannot be answered, and this may provoke a crisis and an eventual change of paradigm. The consequence may be a scientific revolution.

Paradigms provide a worldview, when they change there is a *gestalt* change. This is the way Wittgenstein had thought changes in language-games. To Kuhn a theory is not the true picture of reality itself, it does not prove that facts are really this or that. There are no fixed essences. The world view provided by the new theory that was "born" in the new paradigm is incommensurable with the previous theory. So there is no true or definitive theory proved once and for all, but paradigms provide method and patterns for research that a community of scientists accept and practice. They provide normative functions, not cognitive ones. They are like maps to probe into nature and to give results. Objectivity is in permanent construction.

However, this epistemology, even with characteristics that resemble pragmatism that is holistic and for which the concept of paradigm is central, ends up with some difficulties that Wittgenstein's "critical epistemology" avoids and may "dissolve."

5. Confronting the Contribution of Both Wittgenstein and Kuhn

Kuhn considers that his scope in the study of science history is to study internal issues that have to do with paradigms. External factors like economy or social problems do not interfere with the practicing of normal science and with paradigms that are accepted and practiced by a community of scientists. But neither the appearance nor the uses of a paradigm are restricted to internal aspects of science. Wittgenstein is not concerned with this division. Science cannot be enclosed in a dome. This division is also a fallback in the difficulties to establish the limits of science, its neutrality or not (see, for instance, Carnap).

Only where there is a paradigm there is progress in science, says Kuhn. Human sciences do not have paradigms, so they do not progress. This is another endless discussion. What is progress after all? Atomic bombs brought progress? Psychoanalysis did not? Science depends upon a paradigm and not upon strategic needs, is it independent of politic decisions? Where do new technologies come from?

Wittgenstein's conception of paradigm avoids these dead ends and impasses. Science is an activity that has no definite and neat frontiers. It is not even a privileged kind of knowledge. Its epistemological and logical status depends on uses, on diverse and different needs. In Kuhn's approach of scientific paradigm and progress of science underlies the difficulty of epistemologies that in a certain way privilege science as the only kind of knowledge that uses paradigm and is distinct from all other knowledge. At the same time Kuhn has a sociologic conception of science that is attached to the current paradigm, since it is practiced by a community of scientists that compete and fail, he rejects to take into consideration external factors, social and political pressures over science and over scientists.

If there is progress it has to be over previous theories, but this has a difficulty, it implies that history of science is cumulative, what contradicts Kuhn's claims that there is not a "development-by-accumulation" in the history of science (1996, 2). Another difficulty is: if theories are only instruments, if there is incommensurability, how can we trust them? Kuhn sustains that theories are descriptions or explanations more and more close to reality, but how can scientists test this approximation?

Kuhn distinguishes a science in its early stages of development, where "different men confronting the same rage of phenomena …, describe and interpret them in different ways" (1996, 17), from advanced stages where a paradigm triumphs. The change of paradigm, as it is a kind of scheme through which the world is seen, implies the change in the worldview. This entails epistemological difficulties, there are two different presuppositions, that reality is the same, that it is constituted of pure facts, and the notion that knowledge schemes are the instruments of knowledge. Such sort of idealism is not compatible with Kuhn's claiming that science and theories are kind of instruments and not exact pictures of reality.

These issues do not arise if one considers Wittgenstein's approach of science, particularly paradigm as models, as rules whose changes provoke changes in the way people or a culture use those paradigms. Scientific activities are not separated from interests, habits, contexts, and uses typical of a culture like ours. Method and proceedings, ideas, insights, practices, values, all these factors count, all of them have a role in the system of scientific propositions. To make experiments, to test facts, to observe them and to verify theories—these activities do not have to follow rigid rules of a rigid method. On the contrary, they simply work and are useful in sciences as chemistry or biology. In our life forms, science is an activity that is not fixed by rational patterns and exclusive paradigms.

Notwithstanding Kuhn's concepts are revolutionary, his contribution to epistemology and to the way we view the history of science are notorious.

6. Conclusion

To sum up and conclude, with important differences, both Wittgenstein and Kuhn contribute by means of the concept of paradigm to extend, innovate, and improve the way we see knowledge in general and scientific knowledge in particular. Wittgenstein is with a broader view, a critic epistemology, for the function of paradigms is in connection with grammar and language-games, in every life form. Kuhn is with an epistemology centered on science. Without a paradigm there is no scientific progress, just as without instruments there are no artifacts, no culture.

Works Cited

Kuhn, Thomas. *The Structure of the Scientific Revolutions.* 3rd ed. Chicago, London: Chicago UP, 1996.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton: Princeton UP, 1979.

Wittgenstein, Ludwig. *On Certainty*. Oxford: Blackwell, 1991. <http://budni.by.ru/oncertainty>.

---. *Philosophical Investigations.* 3rd ed. Trans. Gertrude Elizabeth Margaret Anscombe. Oxford: Blackwell, 2001.

Philosophy Study, ISSN 2159-5313
June 2012, Vol. 2, No. 6, 417-431

Resolving Insolubilia: Internal Inconsistency and the Reform of Naïve Set Comprehension

Neil Thompson

This paper seeks to identify the minimal restrictions that need to be placed on the naïve comprehension principle to avoid inconsistency in set theory. Analysis of the logical antinomies shows that at the root of inconsistency in naïve set theory are certain "self contradictory" predicate functions in extensional set descriptions containing the matrix "$\neg(x \in y)$" (or "$\neg(x \in x)$") rather than "size," vicious circularity, or self-reference. A reformed set comprehension system is proposed that excludes extensional set descriptions that conform to the formula, $(\forall x)\ (\exists y)\ (x \in y \Leftrightarrow P\ (x)) \Rightarrow (\exists u)\ (u \in y \Leftrightarrow \neg(u \in y))$, from comprehension and otherwise preserves the ontology of naïve set theory. This reform avoids the paradoxes by scrutiny of a set's description without recourse to type or other constructivist limitations on self-membership and has the most liberal rules for set formation conceivable including self-membership. The intuitive appeal for such an approach is compelling because as a revision of naïve set theory, it allows all possible set descriptions that do not lead to inconsistency.

Keywords: set theory, logical paradoxes, reforming the naïve comprehension principle

1. Prolegomenon

Set theory has long been instrumental to our ideas of formal logic and mathematics, yet no totally satisfying responses have been given to the following concerns:

(1) What is at the root of the logical antinomies?

(2) What set descriptions (or their predicative functions) must be excluded from the range of possible descriptions because they threaten systemic consistency?

(3) If any approaches or limitations to set comprehension are not based on purely logical ideas, how might we justify those answers philosophically?

It is claimed in this paper that:

(1) Most of the logical antinomies are occasioned not by such features as self-reference, vicious circles, or a set's being "oversize" but by a set description containing a matrix of the form "$\neg(x \in y)$." This is most notably to be found in Russell's paradox in the form "$\neg(x \in x)$."

(2) Set descriptions containing such a matrix must be excluded from any set comprehension system claiming to be consistent. Such an exclusion principle also defuses the logical antinomies by blocking the comprehension of sets instrumental in the generation of the inconsistency occasioned by the paradoxes. Indeed it emerges that the system containing the balance of naïve set comprehension after such exclusion must be consistent.

(3) Such exclusions can be justified by purely logical ideas based on limitations around the features of the

Neil Thompson, BA, LLM; main research fields: Logic, the Foundations of Mathematics.

equivalence biconditional "⇔," which is found in all extensional set descriptions.

(4) The deeper philosophical, and indeed epistemological, import of these problems is, of course, immense. If the foundation of mathematics, set theory, is based not on logic but on *ad hoc* constructivist ideas or if we can have no profound conviction that it is consistent, then we are forced back on vague and perhaps unreliable intuitive ideas or even the view that mathematical ideas are based on *a priori* "synthetic" truths to support the foundations of mathematics and its claim to be intrinsically true.

2. Extensional Set Descriptions and Equivalence

A set might be described in at least three ways: *iteratively* by listing its members, *intensionally* by reference to a property applicable to the set treated as an individual; or *extensionally* by reference to a property held in common by its individual members. Extensional descriptions are made through a set descriptive formula stating that an individual's membership of the set and its satisfying the set's identifying predicate function are materially equivalent.

An intensional description of a set that specified that a property be both true and false of that set is itself logically false by the law of non-contradiction. Indeed any such description of an object is logically false on this basis—this is not a principle restricted to sets. As for extensional set descriptions, a predicate specified to be both true and false of an individual must denote the null set.

It may have been originally thought that any predicate designating the membership of a set must simply be true or false (but not both) of any possible member. Such an assumption about extensional descriptions is, as is demonstrated by Russell's paradox, false. It amounts to the naïve set comprehension principle: the idea that for every property there is a unique set of individuals, which have that property: $(\exists y)\ (\forall x)\ (x \in y \Leftrightarrow P(x))$.[1] The predicate $P(x)$ rendered as two contradictory conjuncts in an extensional description thus merely denotes the null set; but when $P(x) = \neg(x \in x)$, a contradiction results when the set so described is itself instantiated for x; or conceivably when "$\neg(x \in x)$" is a conjunct within P so that it can be inferred by detachment. The matrix "$\neg(x \in y)$" in a predicate will be called "the Russell matrix" although in most cases it is the narrower expression "$\neg(x \in x)$" that is to be found in objectionable set descriptions such as that of the Russell set itself.

3. Reform of Naïve Comprehension: The Russell Exclusion

As will become clear, in virtually all the logical antinomies[2] a set is described in the course of their exposition by a property constituted by, or containing, "$\neg(x \in x)$" or "$\neg(x \in y)$." (The scant exceptions do not detract from this general principle and are explainable on the basis that, for example, a set described in the exposition of Burali-Forti's paradox is a logically false intensional description.)

Where that matrix (or derivations from it) is excluded from predicates that might give rise to set comprehension (call this "the Russell exclusion"), what remains of the naïve set comprehension ontology is true on logical grounds because if the right-hand side of the biconditional is equivalent to the negation of the left-hand side, the description as a whole is false, but otherwise an individual is simply a member of a set if the predicative function is true of that individual. Conversely and trivially, the predicative function is not true of an individual if that individual is not a member of the set.

So if we apply the Russell exclusion to set descriptions of the naïve set ontology, what remains of naïve comprehension is: where $P(x)$ is itself a pair of contradictories, or otherwise false of any object, the set described is the null set; otherwise $P(x)$ is either true or false of any individual, so that the description (the biconditional as a whole) is indeed tautologically or at least linguistically true because the set description is a legitimate naming of a collective individual object whose members satisfy the relevant predicate.

Considered a singular proposition a set description can only lead to inconsistency if it is logically false. An extensional set description is only logically false if its predicate function is or contains the Russell matrix, which makes the biconditional false.

Putative extensional set descriptions ruled out by the Russell exclusion are true or false of some individuals but generate contradictions for others. Of course any intensional description constituted by a pair of contradictories is (trivially) false of any individual set. A singular further case is the Goldstein series of sets that is a description of a series of sets or set schema that is inherently contradictory.

So most (but certainly not all) putative set descriptions conforming to the naïve comprehension principle are tautologically true in the sense that such descriptions are truths of language or naming. Set descriptions that are impermissible because of the Russell exclusion imply a contradiction since membership of some individual implies non-membership and any such description is thus itself false. The surviving (permissible) descriptions after the operation of the Russell exclusion, if treated as additional axioms and added to the axioms of propositional logic and the first-order logic, as tautologies preserve logical truth and thus result in a system logically or linguistically true and (ergo) consistent.

It is the internally contradictory, logically false content of certain putative set (or subset) descriptions—rather than the traditionally attributed vices of circularity, self-reference per se, a set's being "oversize," or even random inconsistency—which is at the root of the paradoxes in naïve set theory.

4. The Scope of the Proposed Reform

This proposal for a reformation of the naïve set comprehension principle is conservatively circumscribed: it is concerned only with sets in the context of pure arithmetic and not (for example) with descriptions of collections of natural or non-existent objects, nor with contingent statements.[3] As far as this reform of the naïve comprehension principle is concerned, the naïve set comprehension principle (which applies only to extensional descriptions) is valid except where:

$$(\forall x)\ (\exists y)\ (x \in y \Leftrightarrow P(x)) \Rightarrow (\exists u)\ (u \in y \Leftrightarrow \neg(u \in y)).$$

Proof of u's existence is to be judged by inference *solely* from the putative set description via the standard rules of inference and general logical principles. Hence the existence of a relevant entity u must be inferred from the set description alone and $P(x)$ must be $\neg(x \in y)$ or at least contain it in some form such as a conjunct.[4] An example is where there are multiple quantifiers in the description, it can be inferred that there is one instance where all the variables take the same identity.

In this paper the resultant set comprehension system will be called "Reformed Set Comprehension" or "RSC" whose ontology is itself a proper subset of all set descriptions permitted by the naïve comprehension principle. Crucially, RSC has no non-logical elements, but the limitations on unrestricted naïve set comprehension are drawn from the limitations and properties of the biconditional used in extensional descriptions.[5]

5. Testing Sets in the Paradoxes under the Russell Exclusion

A survey of the known paradoxes inductively supports the RSC approach to workable set comprehension. In Burali-Forti's paradox, the (non-empty) set described is the next largest ordinal to the set of all ordinal numbers. The next highest ordinal is itself an ordinal (because of its own description) and (by definition based on the concept of the ordering of the series of ordinals) larger not only than all ordinals including its immediate predecessor (the set of all ordinals) but also than itself. In this case the set concerned is described intensionally and not extensionally by reference to its membership; rather, it is purportedly described by reference to two inconsistent properties of a specified set.

In Russell's paradox the Russell set, $R = \{x: x \notin x\}$—which is a putative set described extensionally, and if taken as an instance of the variable in its own set description—yields a contradiction. Of course the existence of R is to be inferred from its own description—if the description is valid. So the Russell set description is excluded from RSC comprehension by the Russell exclusion.

It is contended that all set descriptions implicated in the known logical antinomies can be shown intuitively to "contain" their own contradictions since the contradiction is inferred by the set description of an invalid description alone; such contradictions arise from instantiations of individuals identifiable by their descriptions. The excluded set descriptions do not so much lead to the inconsistency of set theoretical systems; rather, they bring to them contradictions already inherent in those descriptions.

The derivation of contradictions from problematic set descriptions represents *reductio ad absurdum* proofs not only that the naïve set comprehension principle[6] is false, but also that some putative "real life" descriptions are false—as Russell's description of the barber on an island of clean-shaven men who shaves everyone, but himself leads us to conclude that there can be no such barber because such a clean-shaven man could never be shaved by anyone else or shave himself.

6. The Inadequacy of Constructivist Hierarchies as a Response to the Antinomies

The standard responses of type and set theory to the logical antinomies are constructivist in inspiration and mandate restrictions on how sets might be members of certain other sets. The solutions employed are directed against self-reference or size which are generally regarded as the principal vices responsible for inconsistency. Self-reference with its perceived self-supporting structures is contrasted with the reassuring solidity of a cumulative hierarchy. Such hierarchical approaches are awkward and prolix in practice.[7] It has never been conclusively demonstrated that size or boundlessness alone is at the root of the derivation of contradictions. Indeed the proscription of useful very large sets (such as a universal set) and wide-ranging generalizations weaken the descriptive range and utility of standard systems and impoverish the imaginative possibilities of set theory ontologies.

These drawbacks suggest that the traditional hierarchical approaches to the problems of set comprehension are merely provisional or approximate solutions to the problems of the logical antinomies and set comprehension. Ideally, as much of the ontology of naïve set theory as possible should be salvaged by a set comprehension principle short of generating inconsistency. Only those sets which give rise to contradiction should be excluded from set comprehension—not every set involving overarching, but apparently useful,[8] generalities. It is worth recalling in this context Cantor's dictum emphasizing the role of consistency in the concept of a set.[9]

7. An Overview of the Significance of These Ideas to the Logical Antinomies

The Burali-Forti paradox involves an intensional description which is an explicit contradiction involving the set of all ordinals. Goldstein's set series (according to its originator) is an explicit contradiction. Cantor's paradox and the paradox of the Greatest Cardinal depend on Cantor's theorem being applicable to "ultimate" sets such as the set of all sets, but its proof in so far as it may be applicable to such a set of ultimate size involves a set description which involves the postulation of a set whose description contains the Russell matrix. Most of the other notable paradoxes involve sets conforming to the Russell matrix in the form "$\neg(x \in x)$."

In Mirimanoff's paradox, the set of well-founded sets as described involves the Russell matrix. Quine's paradox schema is also based on a multi-quantified extension of the Russell matrix.

7.1. The Burali-Forti Paradox

Burali-Forti's paradox is as follows: an ordinal[10] is a set; and each ordinal is the set of all ordinals that precede it. So the set of all ordinals, *On* must exist and it must be an ordinal itself—obviously the largest of all the ordinals. However, since *On* is an ordinal it is suggested that we can form the ordinal *On* + 1. Since *On* + 1 > *On*, we must conclude that *On* both does and does not contain all the ordinals.

However, if *On* is properly described, every ordinal must be a member of it and so any entity, which is not a member of *On*, is not an ordinal. If we attempt to postulate *On* + 1, then the resultant entity must be distinct from the membership of *On* and must yet be an ordinal. So *On* + 1 cannot be an ordinal because it is not a member of *On*. *On* + 1 must be and not be an ordinal and thus its (intensional) description encompasses an explicit contradiction.

More formally (Priest 1993, 28), consider the property: *x* is an ordinal; the set of ordinal numbers, *On* and the function $\delta(x)$, the least ordinal greater than every member of *x* (abbreviation: $logr(x)$). By virtue of its description, $logr(On) \in On$ and $logr(On) \notin On$. There are accordingly two contradictories in the (self-contradictory) intensional set description of $logr(On)$.

To provide a disabbreviated description of $logr(On)$ would be a protracted undertaking, but nonmembership of any predecessor in an ordinal progression is obvious because all ordinals are structured into a rising series with each successive ordinal being the set of all preceding ordinals.

So the successor of any ordinal number is not a member of its predecessor by definition; hence the putative set $logr(On)$ is not a member of *On* by its description.[11] Thus the extended definition of $logr(On)$ requires that both $logr(On) \in On$ and $logr(On) \notin On$. An antinomy along somewhat parallel lines to Burali-Forti's is that propounded by Loïc Colson as "a variant of Burali-Forti's" (2007, 33).

The Colson example arises out of the family of all well-founded sets. A well-founded set is, for such purposes, a pair; the first member of which is itself a set that has no infinite decreasing sequence of elements of that set. The terms *successor* and *family* of well-founded sets are defined and then the sum of a family of sets and all are shown to be well-defined sets.

The term morphism (a function of ordering between two well-founded sets) is then defined and finally the set of all well-founded sets, I_0. The sum of the latter and its successor are considered and it is shown that it is the successor of the sum of I_0. It is then shown that morphism both does and does not exist between the sum of I_0 and its successor. On the one hand, by its description the successor of the sum of I_0 is well-founded. On the other hand, there can also be no morphism from a successor of a set to itself by definition. It is suggested that the result is a contradiction with the successor of the sum of the set of all well-founded sets being from its intensional description both well- and not well-founded.

7.2. Russell's and Similar Paradoxes

Russell's paradox has been discussed. A variant of Russell's paradox is Zermelo's, which was probably discovered by him as early as 1899 (Rang and Thomas 1981, 18). Zermelo's version postulates a set *M* that contains each of its subsets *m*, *m'* ... as elements. Subsets of *M* will also contain certain subsets as elements and some that contain other subsets but not themselves as elements. The set of all subsets of *M*, called M_0, is considered which is readily shown both to contain and not contain itself as an element (Rang and Thomas 1981, 17). In essentials this is the Russell set and is covered by the Russell exclusion.

7.3. Ramsey's Paradox

It has already been explained how the set description of Russell's set is an extensional description, giving rise to a contradiction derived from self-instantiation. Ramsey's paradox also involves a self-instantiation in the area of relations.[12] The description of Ramsey's relation contains a Russell matrix that can be instantiated by the relation itself to produce a contradiction.

7.4. Cantor's Paradox

Cantor's paradox is a paradox involving the sets of ultimate size. Its relationship with the Russell exclusion is not immediately obvious although Quine[13] was aware of its connection with the Russell set. Specifically, it involves a diagonalization out of a set of ultimate size, the set of all sets, *S*.[14] The contradiction is based on the application of Cantor's theorem (the power set of a set is always larger than the original set) to the set of all sets. The contradiction is that if Cantor's theorem applies to the set of all sets then its resultant power set can be shown to contain a member, which is not a member of *S* and therefore not a set. The usual account of Cantor's theorem firstly shows that the power set of a set cannot be smaller than the original set (*s*) because for every member of *s*, a unit set (a subset) exists containing just that member of *s*—so obviously the members of *s* and their unit sets can be correlated on a one-to-one basis.

Next, it is shown that the power set cannot be the same size as *s*. If it is the same size, then the members of *s* and $P(s)$ can be correlated one-to-one. Let $f(x)$ be the function correlating the members of each of *s* and $P(s)$. Then postulate subset t: $x \in t \Leftrightarrow \neg(x \in f(x))$. But *t* must have a correlate *y*. Is *y* a member of *t*? Taking *y* as an instantiation in *t*'s description, we get $y \in t \Leftrightarrow \neg(y \in f(y))$. But given that $t = f(y)$, it follows that $y \in t \Leftrightarrow \neg(y \in t)$. Thus a contradiction is generated, but *t*'s description is covered by the Russell exclusion so the proof is blocked.

An alternative formulation of Cantor's paradox is to be found in Narayan Raja (2005), who draws on ideas from Yablo's paradox. This proof (if it were applicable to *S*) and its power set would lead to an alternative formulation of Cantor's paradox. Raja's proof starts with the assumption that there is a one-to-one mapping, *M*, between a set *X* and its power set $P(X)$. Raja postulates the existence of a "trace" which is a recursive sequence of members of *X*, s_0, s_1, s_3 ... s_j which starts with an arbitrary member of *X*, s_0, and in which each subsequent member of the sequence is a member of *X* but is also a member of a mapped counterpart in *M* belonging to $P(X)$, $M(s_{j-1})$ of the subset mapped by *M* to its predecessor s_{j-1}. The result amounts to a mapping of certain of the members of *X* and $P(X)$. But the effect of this "mapping" is made disjoint by the membership of s_0, s_1, s_3 ... s_j being paired by membership with the corresponding counterpart of their predecessor. Raja sees this requiring the postulation of a subset *N* that cannot be part of a one-to-one mapping with a member of *S*. Raja's proof of this also requires the postulation of an element *n* that is postulated as being mapped to *N*. Because it is a subset of *S*, it is a member of $P(S)$ but it is not a member of the set *S*. In the case of *S*—the set of all sets and its power set $P(S)$—there are no examples of elements that cannot be part of a possible trace that is not terminating. Take any member t and assume that it is mapped to its unit set $\{t\}$. A possible trace satisfying Raja's criteria is "t, t, t ..." which is non-terminating yet satisfies Raja's criteria. We can imagine it being mapped to "$\{t\}$, $\{t\}$, $\{t\}$ ..."; so in such a case each "t" after the first is mapped to a unit set of itself. This means that membership of N is either covered by the Russell exclusion or the description of N is a logically false intensional description where S and $P(S)$ are concerned. Fundamentally, no subset can be a member of $P(S)$ and not be a member of S. Since N is excluded, so is n, which is described by reference to it. Accordingly, Raja's proof has no application to S and $P(S)$ under RSC.

More recently Raja (2009, 2174) has offered one further approach to proof of Cantor's theorem suggesting that the set N is "negation-free" but the point for current purposes is not whether N is negation-free but whether its description can be established under RSC in relation to S and $P(S)$.

The principle of Cantor's theorem, applicable generally to sets, is thus not applicable to the set of all sets, S, under RSC. A similar analysis is applicable to Hilbert's attempt[15] to construct an antinomy along parallel lines to Cantor's paradox. The difficulty in measuring the relative magnitude of the set of all sets and the set of all its subsets is to be found in our efforts to establish that there is a set which is a subset but is yet not a set.

Quine pointed out the problems with a proof in relation to sets of ultimate size[16] in discussing his "New Foundations," even if his precise results cannot be repeated in a comprehension system (like RSC) that allows unit sets. The set t becomes the set of all sets that are not members of themselves in "New Foundations."

Though Cantor's theorem under RSC has no application to S, it is not thereby established that the power set of S is the same size or smaller than S. To do so would seemingly require us to nominate a set or subset which was not a member of S. It may simply be that the standard concepts of cardinality or ordinality (including bijection) have no standard application to sets of ultimate size such as S and $P(S)$. Something of a very broadly similar kind happens in Neumann-Bernays-Gödel set theory as well where "outsize" sets are treated differently from ordinary sets.

7.5. The Paradox of the Greatest Cardinal

This paradox can be regarded as a variation of Cantor's (Quine 1951, 128-29). The set of all sets, S, must be regarded as the set of all its subsets; so S also must be regarded as its own power set. But Cantor's theorem shows that the cardinal of a power set must be larger than the cardinal of the set itself. However, it has already been seen that Cantor's theorem is not applicable to S under RSC.

7.6. Mirimanoff's Paradox

Mirimanoff's paradox is concerned with the cumulative hierarchy. Call a set x, non-well-founded, if there are $x_0, x_1, x_2, \ldots x_n$ such that:

$$x \in N \Leftrightarrow \ldots (x_3 \in x_2) \;\&\; (x_2 \in x_1),$$

where N is the set of all non-well-founded sets.

That is, a non-well-founded set has an infinitely descending regressive membership sequence; but the variables do not need to be distinct, so if $x \in x$ then x is non-well-founded because:

$$\ldots (x \in x) \;\&\; (x \in x) \;\&\; (x \in x).$$

This can be expressed as "... $x \in x \in x \in x$" which follows in the cases where $\ldots x_2 = x, x_1 = x, x_0 = x$.

Hence, by definition, there is a subset of the set of non-well-founded sets that are members of themselves.

Well-founded sets are those that are not non-well-founded. Hence, it follows that (equally by the definition of the set of well-founded sets) there must be a subset of the set of well-founded sets that is the set of all sets which are not members of themselves.

Membership of W, the set of well-founded sets, is:

$$x \in W \Leftrightarrow x \in N,$$

which gives us for some subset, w_I, of W:

$$x \in w_I \Leftrightarrow (x \notin x) \,\&\, (x \notin x) \,\&\, \ldots,$$

which obviously contains a Russell matrix. This subset is the Russell set and its description is excluded under RSC. As far as the set description of W as a whole is concerned, a contradiction ensues. Since all its members are well-founded, it too is well-founded. But that means that $W \in W$ itself gives rise to the infinite regress … $W \in W \in W \in W$. So it cannot be a member of itself if it is a member of itself.

7.7. Quine's Rising Series of Paradoxes

Quine explained how an infinite series of self-contradictory set descriptions can be generated (as it turns out, by using the Russell matrix) to generate self-contradictory set descriptions: "The matrix '$\neg(x \in x)$' is the first and simplest of an infinite series of matrices," viz:

$$\neg(x \in x),\ (\forall y)\ \neg(x \in y \,\&\, y \in x),\ (\forall y)\ (\forall z)\ \neg(x \in y \,\&\, y \in z \,\&\, z \in x),\ \ldots.$$

All of these share the same peculiarity. In the case of each of these matrices as seen in the case of "$\neg(x \in x)$," the assumption of "a class of all entities x such that …" leads to a contradiction. There is no class w such that:

$$(\forall x)\ (x \in w. \Leftrightarrow \neg(x \in x)),$$

nor any such that:

$$(\forall x)\ (x \in w. \Leftrightarrow (\forall y)\ \neg(x \in y \,\&\, y \in x)),$$

nor any such that:

$$(\forall x)\ (x \in w. \Leftrightarrow (\forall y)\ (\forall z)\ \neg(x \in y \,\&\, y \in z \,\&\, z \in x)),$$

and so on, as Quine establishes by a metatheorem in his system. Under reformed set comprehension, all such descriptions are excluded because they can be resolved by instantiations to conjuncts of the Russell matrix.

7.8. Goldstein's Paradox

Goldstein's paradox is controversial.[17] It does not appear to be based on the set description of a particular set. Rather it involves the postulation of an infinite series, K, of sets such that an object is a member of any given set in K if and only if it is not a member of any subsequent set in K. Each set in the sequence is defined as follows:

D_1 $\quad \forall x(x \in C_1) \Leftrightarrow$ for all $k > 1$, $x \notin C_k$,
D_2 $\quad \forall x(x \in C_2) \Leftrightarrow$ for all $k > 2$, $x \notin C_k$,
…
D_n $\quad \forall x(x \in C_n) \Leftrightarrow$ for all $k > n$, $x \notin C_k$.

Its formulation proceeds "by taking some arbitrary object $\propto$." We can see that it satisfies the membership requirements for C_n *by* D_n: $\propto \in C_n \Leftrightarrow$ for all $k > n$, $\propto \notin C_k$.

Now if $\propto \in C_n$,
then $\propto \in C_{n+1}$,
and, indeed, for all $k > +1$,
$\propto \notin C_k$.
From $\propto \notin C_k$,
and D_{n+1},
$\propto \notin C_{n+1}$,
which contradicts $\propto \in C_{n+1}$ (above).
By *reductio*, $\propto \notin C_n$.

Goldstein points out that from a similar process of assuming $\propto \in C_{n+1}$, it can be shown that:

$\propto \notin C_{n+1}$,
and similarly,
$\propto \notin C_{n+2}$,
and so on so that for all $k > n$,
$\propto \notin C_k$.
And from this and D_n, it follows that:
$\propto \notin C_n$.

Goldstein's paradox thus describes sequences that are sets of sets in which every member of the subsets forming the infinite sequence consists of those elements which are not members of any subsequent set in the sequence. Goldstein diagnosed the vice in the sequence as being in the format of D_n: "If, by some D_n, some object is in C_n, then that object is not in C_{n+1}, therefore by D_{n-1} the object is in C_{n-1}, therefore it is not in C_n. So D_n (with the connivance of D_{n-1}) is in effect stipulating that a certain object is both in and not in C_n" (1994, 225).

Goldstein concludes that C_n in all its forms should be rejected as circular and/or self-contradictory. It is clear that C_n is in itself a set of sets described by reference to another set of sets—its successors in a sequence. The definition schema requires that every set in the D_n sequence be defined by reference to its successors in the sequence.

So the description of every set C_n requires that a member of that set be both a member and a non-member of C_{n+1}. Thus if Goldstein's schema could be shown to have an occurrence which generated a set description, then the (intensional) description of that set would be overtly self-contradictory; such a description would be false.

7.9. Inconsistency Occasioned by Set Descriptions Involving the Law of Identity

The Russell exclusion is applicable to certain descriptions where the right-hand side of the "$\Leftrightarrow$" in the set description formula generates the Russell matrix because of some instantiation giving rise to the law of identity implying $\neg(x \in s)$.

Take the following example of a set description containing the "$\Rightarrow$" connective:

$x \in s \Leftrightarrow \forall y\, \forall z\, (\{y\} = \{z\} \Rightarrow \neg(x \in z))$. (Consider the instantiation when $y = z = s$.)

Similarly for:

$x \in s \Leftrightarrow \forall y\, \forall z\, ((x \in s) \Rightarrow \neg(\{y\} = \{z\}))$,
and $x \in s \Leftrightarrow \forall y\, \forall z\, (\neg(\{y\} = \{z\}) \vee \neg(x \in z))$,

which all lead to the contradiction $x \in s \Leftrightarrow \neg(x \in s)$ where $y = z = s$.

All are covered by the Russell exclusion. Another covered is:

$$x \in s \Leftrightarrow \forall u\, \forall v\, \forall y\, \forall z\, ((\{y\} = \{z\}) \Rightarrow \neg(\{u\} = \{v\}))\ \&\ ((x \in s) \Rightarrow \neg(\{u\} = \{v\})),$$

which generates a contradiction when $u = v = y = z = s$.

These further set descriptions excluded under RSC are not paradoxical in character, but they do show that the well-known antinomies are not alone in giving rise to inconsistency from set descriptions.

8. Reformed Set Comprehension and Other Logical Systems

RSC is intended to capture the set descriptions implicated in all the standard paradoxes, but permits all other set descriptions that do not come within the exclusionary schema. It is designed only for use with the standard rules of inference of classical logic and set theory, although its abstraction schema does not exclude non-well-founded set descriptions in general. No claims are made here about its general application to other logical systems or approaches.

9. Establishing the Invalidity of a Set Description under RSC

Most importantly, a set description could be scrutinized under RSC for invalidity by looking at its propositional structure or the propositional structure of relevant "internal" instantiations. Disabbreviation may be required. This is a relatively simple undertaking of a simple truth tabular kind because the number of individuals implicated by any description in most cases will be relatively small.

One somewhat counterintuitive result remains under RSC: it might be thought that "under certain circumstances a definable collection ['*Menge*'] does not form a totality" (ed., van Heijenoort 1967, 125), as Russell[18] observed in correspondence to Frege. Under RSC not all concepts describe as set. So, for example, although non-self-membership is not a legitimate criterion in RSC to describe a set, it is a perfectively tenable concept or predicate in other respects.

10. Application of the Principles of Such Analysis to the Epistemic Paradoxes

The possible relevance of these principles to the epistemic paradoxes is fairly obvious but is being only touched on here. In the case of a common form of the Liar, in which a Cretan claims that all Cretans are liars, such statement implicitly claims itself to be true (because of the material equivalence of the liar sentence and a sentence claiming it to be true) and the set of false statements because it is a statement made by a Cretan which according to the statement itself is thus untrue. To generate contradictions from such statements, further non-logical assumptions are required using biconditionals usually called truth schemas or T-schemas (Priest 1994, 30).[19] It is the truth biconditional that is at the heart of the generation of the inconsistency because without limits being placed on T-schemas the Liar sentence allows a sentence to implicitly say of itself that it is false.

11. Circular Definitions and Descriptions

There is another objection to any set comprehension system that allows self-membership: the intuitive objection of circularity. Might allowing self-membership of a set be viciously "circular"? In traditional logic, circular definitions of terms are illegitimate. A circular definition was said to be one in which the term to be defined occurs within the defining expression in such a way that no one could understand the *definiens* who did not already understand the *definiendum*.[20] (The somewhat obscure qualification to the otherwise purely formalistic features of the concept of circularity is emphasized.) Similar principles would apply to descriptions that are designed to uniquely designate individuals. The fact that a definition is circular does not, of course, mean that the validity of the *definiendum* is completely precluded. The individual's existence or a term's meaning can be salvaged if it can be established by an alternative (non-circular) definition. Circularity is not expected to generate inconsistency—merely a deficit in meaning or an unclear reference to an individual.

The constructivist conception of a set offers instructive insights into sets in finite universes. Russell's claim that propositions about classes can be interpreted (iteratively) as propositions about their elements becomes "literally true" in the case of a finite group of individuals as Gödel (1944, 144; 134 of reprint) observed since "x_m" is equivalent to:

"$x = a_1 \vee x = a_2 \vee \ldots \vee x = a_k$,"

where the a_1's are the elements of m and further that "there exists a class that …" is equivalent to "there exist individuals such that...." Yet it is disingenuous to think it is possible to dispense with generalization in this way either in practice or theory. Of course on this basis, self-membership would result in circularity of definition—the universal set would have to appear in the *definiens* and *definiendum* on both sides of its membership description constructed in this iterative fashion. The usual form of description of a universal set using the predicate expression of identity is certainly not circular in this way. It offers just as much an argument *against* the iterative concept of a set as a justification for it. At best circularity in the context of set descriptions suggests a lack of coherent meaning or identity, not inconsistency.

Identifying precisely how circularity might arise in connection with the description of sets in a formal system is not clear. The most basic approach to circularity is to approach it in a narrow formal sense and ignore the qualification. It is enough that if the *definiendum* appears as an expression in the *definiens*, the putative definition is circular. The Russell set is not described circularly in this narrow sense since it is constructed from a property, which is itself a construct of a variable, negation, and set membership. Nor do the other logical paradoxes offer circular definitions or descriptions in this standard formal sense.

Despite their reductionist view of definition, logicians such as Russell[21] have long claimed (and many must have intuitively suspected) that there is some kind of circularity involved in the logical paradoxes—although no one appears to actually claim that circularity in the traditional sense is at issue; nor even in the sense discussed above of Gödel's suggestion about how propositions about finite classes can be treated as propositions about their elements.

Russell's attempts to diagnose the problem varied in detail. In "On Some Difficulties in the Theory of Transfinite Numbers and Order Types" he wrote, "Given a property φ and a function δ, such that, if φ belongs to all members of u, $\delta(u)$ always exists, has the property φ, and is not a member of u; then the supposition that there is a class w of all terms having property φ and that (w) exists leads to the conclusion that $\delta(x)$ has and has not the property φ" (1906, 142 of reprint). Elsewhere he emphasized a perceived ambiguity in the use of the quantifier "all" (Russell 1908, 154; Whitehead and Russell 1910-13, 37).

Type theory draws on the vicious circle principle or circularity in general only as a generalized philosophical justification; it is not an axiom or a theorem of type theory. The principle was subjected to potent attack from Gödel (1944, 135)[22] and has latterly been undermined by claims that Yablo's (1993)[23] paradoxes can be devised that do not depend on self-reference. Others have questioned circularity as the source of the paradoxes (Barwise and Moss 1996, 60).[24]

12. Does Self-Membership of Sets Necessarily Involve a Vicious Circularity?

It is nowhere made clear why self-membership must necessarily lead to contradiction or even why it is objectionable in practice apart from those examples where it is associated with the derivation of a contradiction. Writing "This page does not contain sentences of over 45 words" is self-referential but hardly problematical. No one has attempted to explain why, if self-membered sets do not per se lead to contradiction, they are to be regarded as illegitimate—except that it is pragmatically expedient to uphold prohibitions that protect systems from contradiction, which happen to prohibit self-membership in general. Intuitively, there seems nothing inherently contradictory in, for example, the set of all abstract ideas being itself an abstract idea.

The universal class, V, can be described as the set of everything that is identical to itself. If there are no hierarchical restrictions, V might be a member of itself. V itself satisfies the requirements of its own membership because $V = V$. And the notion of the universal set—loosely, "everything being a part of everything," is not obviously implausible or vicious. Nor does using negation by itself create a vicious circle. George Boolos writes:

> It is important to realize how odd the idea of something's containing itself is. Of course a set can and must include itself (as a subset). But contain itself? Whatever tenuous hold on the concepts of set and member were given one by Cantor's definitions of "set" and one's ordinary understanding of "element", "set", "collection", etc is altogether lost if one is to suppose that some sets are members of themselves. The idea is paradoxical not in the sense that it is contradictory to suppose that some set is a member of itself, for after all, "$(\exists x)$ $(S_x$ & $x \in x)$" is obviously consistent, but that if one understands "$\in$" as meaning "is a member of", it is very peculiar to suppose it is true. (1971, 219-20)

Yet the diffuse idea of peculiarity hardly amounts to a conclusive, logical objection. Further, is "contain" in this context a rigorous conception for sets? A set described extensionally is simply a plurality treated as an individual whose members conform to a predicate function. The approach of RSC is certainly at odds with such thinking about sets and even with Cantor's concern that "the total of everything thinkable" was an "absolutely infinite or inconsistent multiplicity" (Hallett 1984, 166), because it was thought that the grasp of such a totality must exclude the instant of that very thought or perception itself. Such an idea probably underlies both Zermelo-Frankel set theory's and type theory's limitations on set formation (Boolos 1971, 220-24). Understandable wariness about self-membership might certainly justify care, but such objections are not as substantial as might be feared.

To be useful as an analytical tool, a set has to be capable of logical treatment as an individual in sentences. In this sense it is a unity and not a plurality. But any intuitive gathering together of the members or a mental grasp of a set as a collection of individuals is not a critical element of a set's legitimacy. Indeed for any infinite collection or even a finite but very large set, it is a task only a Cantorian divinity could accomplish. As Jaakko Hintikka remarks, such set constructions are purely notional, "they are not constructions which anyone is supposed to be able to carry out" (1996, 154). We are quite capable of deciding, for example, whether an individual is a soldier without gathering together every soldier in the world in our mind.

Any conception of a set requiring that the members must be able to be mentally gathered together or figuratively "grasped" is logically otiose. It is the ability to generalize and process generalizations that makes logic (and sets) powerful analytical tools through concepts, not clumsy and unmanageable theoretical iterations of indefinite or infinite length. It is neither necessary to be acquainted with all the individuals in a set nor to have some cognition of them as an assembled collection in order to reason or generalize about its membership. The key is whether a particular individual satisfies the requirements of membership. A set which is within the membership range of its own description is not for that reason alone unintelligible, counterintuitive, or illegitimately circular. Yet perhaps the term "collection" in the sense of a plurality considered as an individual best represents the underlying logical notion of a set better than any other term available. A collection is a plurality of individuals conforming to some property itself treated as an individual. In order to speak of its constituents collectively, it is not necessary to know each of them individually or gather all of them together physically or imaginatively.

The role of circularity in the logical antinomies is merely illustrative and intuitive of their self-contradiction, but offers no satisfying analysis of, or solution to, them.

13. Other Set Theory Axioms and RSC

Because of the absence of hierarchical limitations on set formation in RSC, its ontology is not finitary (successive iterations of unit sets of the empty sets give us an infinite collection, for example) so no axiom of infinity is necessary. The axiom of choice is non-logical in its content. It requires the postulation of not a set but a particular predicate function that may give rise to a set. The axiom of choice thus appears to be optional for a set theory employing RSC.

14. Conclusion

The constructivist responses to the logical antinomies in standard set theory represent *ad hoc* "fixes" of the problems of set comprehension and inconsistency and lack a satisfying philosophical rationale. This paper proposes a purely logical response to set comprehension that permits self-membership but can still claim to be consistent. The key to defusing the logical antinomies is recognizing that in certain cases the predicate function of an extensional description (in a form—the Russell matrix—the archetype of which is found in Russell's paradox) is logically false for certain instantiations as the left-hand side of the biconditional becomes the negation of the right-hand side. All other naïve set theory extensional descriptions are simply true or false of any possible individual member. It is this feature of set descriptions, which leads to contradiction, rather than the conventional diagnoses of self-reference, "size," or circularity. The proposed reformed set comprehension principle excludes such self-contradictory set descriptions so the resultant system has especially liberal rules for set formation including self-membership. The intuitive appeal of such an approach is compelling. Above all, it does not rely on non-logical elements in its conceptual approach or justification and can claim to be consistent on the basis of the principles of ordinary logic.

Notes

1. Attributed to Cantor but perhaps owing more to Russell or (in a different form, the *Basic Law V*) to Frege.

2. The term "antinomy" is strictly probably preferable to the term "paradox" (as Zermelo explained in 1907) because "paradox" contains nothing of the sense of inner contradiction found in the paradoxes of Russell and Burali-Forti (Peckhaus and Kahle 2002, 157). This paper makes no substantial distinction between them although Zermelo's notion of antinomy is strictly more apt to the revised comprehension system proposed here.

3. Some Curry set paradoxes are not expressible in the limited logical system in contemplation here because it does not include the notion of contingency. A set theoretical version of Curry's paradox is: Consider the sets $\{x: x \in x \Longrightarrow P\}$ and $\{x: x \in x \Longrightarrow \neg P\}$ where P is a contingent statement which could be true or false. Assume that it is a valid description and call the set so described C. So instantiating C in its description we get $C \in C \Leftrightarrow (C \in C \Leftrightarrow P)$, and accordingly, P follows. However, it might be contended that a set theory might equally tell us that the following set description is valid: $\{x: x \in x \Longrightarrow \neg P\}$. Call this set "$D$." Instantiating D in that description we get $D \in D \Leftrightarrow (D \in D \Longrightarrow \neg P)$, and similarly, $\neg P$ follows. However, there is no compelling

reason why such contingent statements as P or $\neg P$ need to form part of a system of set theory anymore than they would be part of *Zermelo-Frankel set theory* or type theory. If a set theory system does not countenance contingent statements, no "Curry" contradiction arises. Any system that did countenance contingency or some comparable notion of modality would require rules for their deployment, one of which must ensure that there is no prospect of both P and $\neg P$ being contingently true under the same circumstances.

4. Or its equivalent: any proposition containing other propositional connectives can of course be converted to one containing conjunction and negation alone.

5. It contains no extralogical elements such as the axiom of reducibility or the axiom of subset formation. A persuasive criticism of axiomatic set theory in this respect is to be found in Hintikka, "Axiomatic Set Theory Frankelstein's Monster?" (1996).

6. Hallett challenges the attribution of the naïve comprehension principle to Cantor: "Nothing like the comprehension principle of the so-called naive set theory follows from Cantor's statements. If 'naive set theory' is characterized as set theory based on the comprehension principle, then this goes back not to Cantor but to Russell [1903] which states that any predicate expression, $P(x)$, which contains x as a free variable, will determine a set whose members are exactly those objects which satisfy $P(x)$" (1984, 38).

7. A persuasive criticism of axiomatic set theory in this respect is to be found in chapter 8 of "Axiomatic Set Theory Frankelstein's Monster?" (Hintikka 1996).

8. Priest (2002, 138) points out that since no variable can range over all propositional functions under Russell and Whitehead's type theory, general principles such as the Law of Excluded Middle, the Axiom of Reducibility, and the Vicious Circle Principle itself cannot be expressed; indeed all such principles are meaningless.

9. See Hallett, 1984. He gives a translation of part of Cantor's 1899 letter to Richard Dedekind as follows: "When … the totality of elements of a multiplicity can be thought without contradiction as 'being together' so that their collection into 'one thing' is possible, I call it a consistent multiplicity or a set" (34).

10. A set S is an ordinal if and only if S is totally ordered with respect to set containment and every element of S is also a subset of S. Any ordinal is the set of all preceding ordinals. A consequence is that every ordinal S is a set having as elements precisely the ordinals smaller than S, so no ordinal can be a member of itself. If there is a set of all ordinals, then that set could not be an ordinal in this sense because it would then be a member of itself.

11. A number is less than another when one of the subsets of the latter is equipollent with the former, which means there must be a member of the latter remaining when the other elements of the numbers are put into a one-to-one correspondence.

12. Let R be the following relation: $<y, z> \in r_x \Leftrightarrow <y, z> \in x$ & $<y, z> \notin y$. If $r_x \notin x$ then $<r_x, r_x> \in r_x \Leftrightarrow <r_x, r_x> \notin r_x$. It is a straightforward self-instantiation generating a contradiction analogously to the Russell set and containing a Russell matrix in the area of relations.

13. See Quine (1937) discussed below.

14. Cantor is said to have proved that the totality of alephs does not exist because if there was such a set, a certain larger aleph must then itself exist based on diagonalization reasoning (see Peckhaus 2002, 29, 157). However, this argument assumes that such an ultimate set can be measured or sized in a conventional fashion and thus shown to be larger than any aleph; but by postulating that the set of all alephs is a super-aleph we are suggesting that some aleph does not conform to the rules (as it were) of alephhood and thus raising a contradictory intensional description. Hilbert's paradox is in substance a variant of Cantor's (Peckhaus and Kahle 2002, 167).

15. See Peckhaus and Kahle, 2002, 165.

16.Cf. Quine: "Since everything belongs to V, all subclasses of V can be correlated with members of V, namely themselves. In view of Cantor's proof that the subclasses of a set k cannot all be correlated with members of k, one might hope to derive a contradiction. It is not clear, however, that this can be done. Cantor's *reductio ad absurdum* of such a correlation consists of forming the class h of all those members of the original class k, which do not belong to the subclasses to which they are correlated, and then observing that the subclass h of k has no correlate. Since in the present case k is V and the correlate of a subclass is that subclass itself, the class h becomes the class of all those subclasses of V which do not belong to themselves. But R3' provides no such subclass h. Indeed h would be $\hat{y}\ \neg(y \in y)$, whose existence is disproved by Russell's paradox" (1937, 91-92, Note).

17. See Goldstein, 1994. But cf. Priest (1997) who challenges whether a set theoretical paradox has been established at all and argues that what is postulated is a set schema rather than a set.

18. More accurately, it might be said that every definable concept does not form a totality (or "collection"—which term is to be preferred).

19. A T-schema relevantly provides that *assertion of a statement* and *a statement that such statement is true* are materially equivalent.

20. See, e.g., Barker, 1965, 201. "A circular definition smuggles the *definiendum* into the *definiens* in such wise as to prevent expansion into primitive notation" (Quine 1961, 242).

21. Poincaré and Weyl also disapproved of "impredicative" definitions and vicious circles. See Church, 1956, 347.

22. Cf. Priest, 2002, 136.

23. Goldstein's paradox has already been discussed.

24. Fraenkel, Bar-Hillel, and Lévy considered and criticized four arguments intended to justify the Vicious Circle Principle: "first it avoids the semantic paradoxes (but there are other ways of doing so); second it rules out self-reference (but self-reference by itself is unobjectionable, as witness the highly useful self-reference in Gödel's arithmetization of syntax); third it avoids the infinite regression entailed in testing the impredicativity of impredicative concepts (but that is not really entailed, for finite proof

procedures often suffice to determine the applicability of impredicative concepts). Only the fourth argument seems plausible: it refers to the nonconstructive character of impredicatively introduced objects" (1958, 176). See also Copi, 1971, 105.

Works Cited

Barker, Stephen. *The Elements of Logic*. New York: McGraw-Hill, 1965.

Barwise, Jon, and Lawrence Moss. *Vicious Circles or the Mathematics of Non-wellfounded Phenomena*. Stanford: CSLI Publications, 1996.

Boolos, George. "Iteration Again." *Philosophical Topics* 17.2 (1989): 5-21.

---. "The Iterative Concept of a Set." *Journal of Philosophy* 68.8 (1971): 215-31.

Colson, Loïc. "Another Paradox in Naive Set Theory." *Studia Logica* 85.1 (2007): 33-39.

Church, Alonzo. *Introduction to Mathematical Logic* (enlarged edition). Princeton: Princeton UP, 1956.

Copi, Irving. *The Theory of Logical Types*. London: Routledge and Kegan Paul, 1971.

Fraenkel, Abraham, Yehoshua Bar-Hillel, and Azriel Lévy. *Foundations of Set Theory*. Amsterdam: North Holland, 1958.

Gödel, Kurt. "Russell's Mathematical Logic." *The Philosophy of Bertrand Russell*. Ed. Paul Schlipp. New York: Tudor, 1944. 123-53. Rpt. *Kurt Gödel: Collected Works*. Vol. 2. Ed. Solomon Feferman. New York: Oxford UP, 1990. 119-43.

Goldstein, Laurence. "A Yabloesque Paradox in Set Theory." *Analysis* 54.4 (1994): 223-27.

Hallett, Michael. *Cantorian Set Theory and the Limitation of Size*. Oxford: Clarendon, 1984.

Hintikka, Jaakko. "Axiomatic Set Theory Frankelstein's Monster?" *The Principles of Mathematics Revisited*. Cambridge: Cambridge UP, 1996. 163-82.

Peckhaus, Volker. "Paradoxes in Gottingen." *One Hundred Years of Russell's Paradox*. Ed. Godehard Link. New York: Walter de Gruyter, 2004. 501-16.

Peckhaus, Volker, and Reinhard Kahle. "Hilbert's Paradox." *Historia Math* 29.2 (2002): 157-75.

Priest, Graham. *Beyond the Limits of Thought*. 2nd ed. Oxford: Clarendon, 2002.

---. "The Import of Inclosure: Some Comments on Grattan-Guinness." *Mind* 107.428 (1998): 834-40.

---. "The Structure of the Paradoxes of Self-Reference." *Mind* 103.409 (1994): 25-34.

---. "Yablo's Paradox." *Analysis* 57 (1997): 236-42.

Quine, Willard Van Orman. *Mathematical Logic*. Rev. ed. Cambridge, Mass.: Harvard UP, 1951.

---. "New Foundations for Mathematical Logic." *American Mathematical Monthly* 44 (1937): 70-80. Rpr. *From a Logical Point of View:* 9 *Logico-Philosophical Essays*. 2nd ed. Cambridge, Mass.: Harvard UP, 1961. 80-101.

Raja, Narayan. "A Negation-Free Proof of Cantor's Theorem." *Notre Dame Journal of Formal Logic* 46.2 (2005): 231-33.

---. "Yet Another Proof of Cantor's Theorem." *Dimensions of Logical Concepts*. *Coleção CLE* 54 (2009): 209-217. <http://www.tcs.tifr.res.in/~raja>.

Rang, Bernhard, and Wolfgang Thomas. "Zermelo's Discovery of the 'Russell Paradox.'" *Historia Math* 8.1 (1981): 15-22.

Russell, Bertrand. Introduction. *Tractatus Logico-Philosophicus*. By Ludwig Wittgenstein. London: Routledge and Kegan Paul, 1963. ix-xxii.

---. "Les paradoxes de la logique." *Revue de Métaphysique et de la Morale* 14 (1906): 627-50.

---. "Letter to Gottlob Frege." 16 June 1902. Rpt. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1933*. Ed. Jean Van Heijenoort. Cambridge, Mass.: Harvard UP, 1967. 125.

---. "Mathematical Logic as Based on the Theory of Types." *American Journal of Mathematics* 30.3 (1908): 222-62. Rpt. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1933*. Ed. Jean van Heijenoort. Cambridge, Mass.: Harvard UP, 1967. 150-82.

---. "On 'Insolubilia' and Their Solution by Symbolic Logic." Rpt. *Essays in Analysis*. Ed. Douglas Lackey. London: Allen and Unwin, 1973. 190-214.

---. "On Some Difficulties in the Theory of Transfinite Numbers and Order Types." *Proceedings of the London Mathematical Society* 2.4 (1906): 29-53. Rpt. *Essays in Analysis*. Ed. Douglas Lackey. London: Allen and Unwin, 1973. 135-64.

---. *The Principles of Mathematics*. Cambridge: Cambridge UP, 1903. 2nd ed. (With new introduction). London: George Allen and Unwin, 1937.

Whitehead, Alfred North, and Bertrand Russell. *Principia Mathematica*. 1910-13. Vol. 1. 2nd ed. Cambridge: Cambridge UP, 1925.

Yablo, Stephen. "Paradox without Self-Reference." *Analysis* 53.4 (1993): 251-52.

Philosophy Study, ISSN 2159-5313
June 2012, Vol. 2, No. 6, 432-438

When Sensory Substitution Devices Strike Back: An Interactive Training Paradigm

Zachary Reynolds
Gordon College

Brian Glenney
Gordon College

A sensory substitution device (SSD) is a technology that translates information for one sensory modality, like vision, into information for use by another, like touch. Though SSDs have been in existence for over four decades, effective training techniques for their use are rarely discussed. In this paper, we compare three training strategies on a particular SSD known as the vOICe. These comparisons were conducted using a minimal but active search and localization task of luminescent discs. These studies show that an interactive training paradigm, which combines the efforts of two trainees at the same time in a tag-like game, is more effective than passive training at a computer console or active training involving search and localization of luminescent discs. This finding supports philosophical views of perception that take interaction with the environment seriously. In particular, we argue that these findings are in support of certain features of the extended mind view as proposed by Andy Clark (2008). In particular, the study suggests that when environmental conditions are responsive to one's behavioral activity, the activity itself is enhanced, as is predicted by the extended mind view.

Keywords: sensory substitution device, vOICe, extended mind, perceptual training

1. Introduction

The vOICe is a sensory substitution device (SSD), a technology that translates visual patterns from a video camera to electro-stimulation patterns on the tongue (Meijer 1992). For instance, when the subject wearing the device points their head camera at a square drawing, they feel a square impression on their tongue. Recent experiments using this device have focused mainly on the experience of "seeing" with the device (Auvray, Hanneton, and O'Regan 2007), or on its use in passive tasks such as object recognition (Chebat et al. 2007), or on very minimal search tasks such as locating individual LEDs in a one-dimensional array (Proulx et al. 2008). Few experiments, however, explore SSD use and the role of interaction with an active environment,[1] a feature emphasized in recent philosophical literature. Two such views, enactivism and extended mind, are worth noting as they suggest distinctive kinds of such interactive conditions (Noë 2006; Clark 2008).

Alva Noë, a proponent of enactivism, writes, "The world makes itself available to the perceiver through physical movement and interaction" (2006, 1). Noë claims that perceptual experience is dependent on our activity within our environment. For enactivism, using one's body in a perceptual task is more critical than using one's brain. Enactivism suggests that use of the device will be best enabled by full-body interaction with

---

Zachary Reynolds, BA, Research Intern, Philosophical Psychology Lab, Gordon College, USA; main research fields: Sensory Perception, Vision.
Brian Glenney, Ph.D., Professor of Philosophy, Department of Philosophy, Gordon College, USA; main research fields: Philosophical Psychology, Philosophy of Perception.

one's environment. The device user will have to engage in an *active* task in order to have the best understanding of how to use it.

Another view that emphasizes active participation is Andy Clark's "extended mind" theory. Clark argues that the perceiver and the world *integrate*, rather than merely interact. Clark writes, "Think of a dancer, whose bodily orientation is continuously affecting and being affected by her neural states, and whose movements are also influencing those of her partner, to whom she is continuously responding" (2008, 24). For the extended mind, the environment is part of your body and brain, integrated by the environment's give-and-take affects on the body's activity. This suggests that in addition to bodily activity, recursive influence on the part of the environment will aid SSD learning. The device user will have to engage in an *interactive* task, where the environment responds to the activity of the user, in order to best utilize it. The success of these two types of training—active and interactive—can be put to the same test to determine which is most effective.

Aside from these theoretical suggestions, it is also reasonable to ask whether the best training regimes are those specialized to train for the skill being tested. How extensive must the regimes be: do they require hours upon hours of training? Lastly, with respect to the particular device in question: might the way one uses the vOICe be distinctive given the different variables available in its software? For instance, the vOICe uses a variable scanning "soundscape," which in normal conditions scans one scene per second, but can increase in the normal range to eight scans per second. Might object recognition benefit from a slower scan rate while increased rates bring more benefit to interactive tasks? This study tests only a small set of these testing variables, focusing on the role of training regime.

2. Experiment

Our preliminary investigation of these three issues—different uses of the device, different training regimes, and different software settings—compares the performance of subjects using the vOICe in a full-body interactive localizing task across three minimal (< 40 mins.) training regimes, all at the highest normal refresh rate of eight scans per second, and all with blindfolded sighted subjects.

2.1. Minimal Localization Task

The minimal localizing task (MLT) required subjects to turn off three hand size luminescent discs that had to be found only by use of the vOICe. Each disc was randomly hung on three walls at varying heights in a 2 × 2-meter dark room. Subjects wore a webcam attached to blind goggles connected to a computer worn in a backpack. The computer transformed the visual images using the vOICe program into a soundscape. In this study 16 total participants, four in each of four different training paradigms, were placed on a "start" line and then told to turn off the three luminescent discs in the room. This task had three timed trial runs per subject.

2.2. Training Groups

The "non-trained" group subjects were given the most minimal oral explanation on how to interpret the sounds from the vOICe. Subjects were only told that the sounds heard in the headphones were related to the visual images captured by the camera. Then, after a brief explanation of their task, they immediately engaged in the MLT.

The "passive trained" group were trained for approximately 40 minutes playing tic-tac-toe against a computer, a training option available on the vOICe program. Subjects passively trained did not wear blindfolds and could open their eyes to look at the screen when they lost the cursor's location. In addition, the scanning refresh rate, adjustable on the vOICe, doubled every 10 minutes. Subjects started the tic-tac-toe game at one scan per second, moving to eight times per second by the half-hour mark. After the training, subjects donned the mobile vOICe device—webcam, blind goggles, backpacked computer, and headphones—and performed the minimal localization task.

The third “active trained” group were given the most minimal explanation used for the “non-trained” group and then immediately equipped with the mobile vOICe device at eight scans per second. They were then asked to turn off three randomly placed luminescent circles by touch while inside a circle of chairs approximately five meters in diameter in a dark room. Subjects had three trial runs, with each run taking between approximately 1 to 10 minutes, averaging approximately 10 minutes of training across the three trials. They took the MLT immediately upon completing this task.

The fourth “interactive trained” group required paired training in a large 20 × 8-meter darkened theatre. Two subjects, each wore the mobile vOICe device in a backpack on the subject’s front side. A luminescent disc was attached to the backpack. Both subjects were asked to turn off their opponent’s luminescent disc before their opponent turned off their own. The training round ended when a subject succeeded at turning off the opponent’s disc. The SSD tag interactive training consisted of five trial games, each lasting between 30 seconds to 2 minutes, with an average of one minute per game. Subjects averaged approximately five minutes of training and then immediately took the MLT task.[2]

2.3. Results

Those interactively trained showed *significantly quicker* completion time of the MLT than subjects without training (Student’s t-test, $t(16) = 3.59$, $p < 0.0025$), with passive training (Student’s t-test, $t(16) = 5.48$, $p < 0.0001$), and with active training (Student’s t-test, $t(16) = 2.25$, $p < 0.039$). (See fig. 1 for the average scores.)

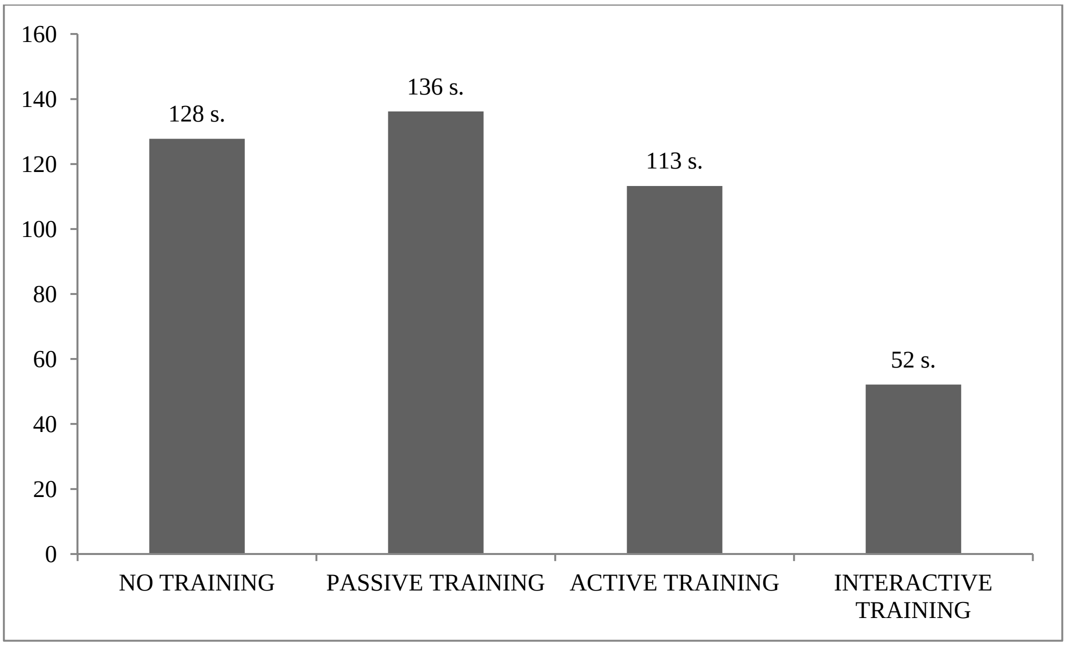

Fig. 1. MLT Completion Averages.

Amongst the three trials themselves, there was an (expected) increase in completion time from the first trial to the final trial due to practice effect, but this was only significant for those without training (Student’s t-test, $p < 0.035$). Those with interactive training showed a significant inverse of this expectation, getting worse with each trial (Student's t-test, $p < 0.024$). (See fig. 2.)

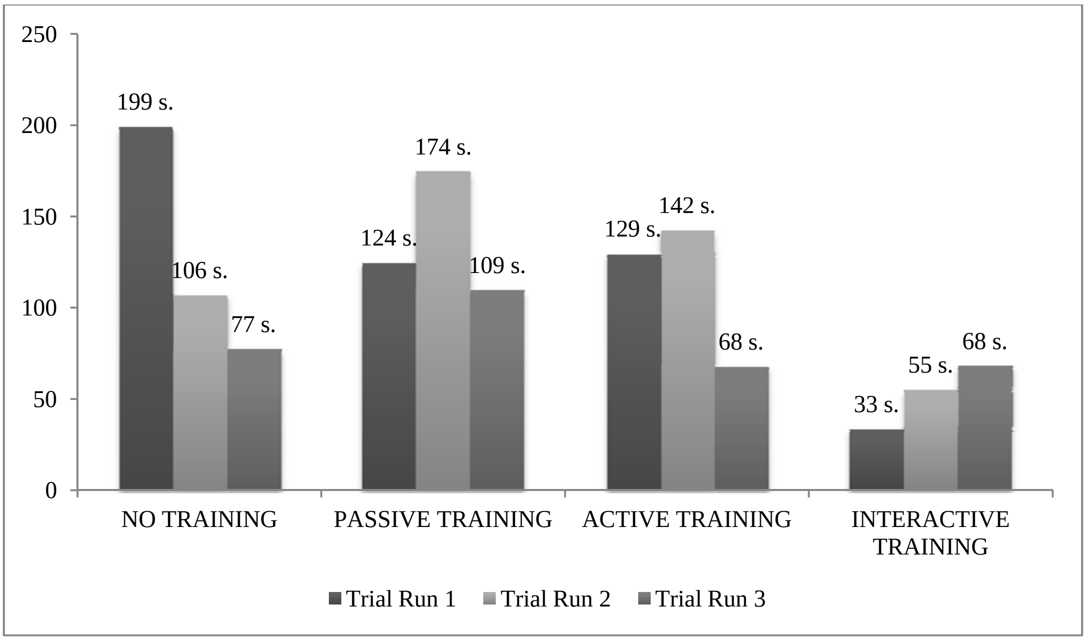

Fig. 2. MLT Averages by Trial Runs.

To emphasize this inversion of the practice effect for the interactively trained, fig. 3 shows that, across the board, individual subjects generally grew slower at the MLT per trial run. This is in contrast to the faster completion times across trials for each subject in the non-trained group.

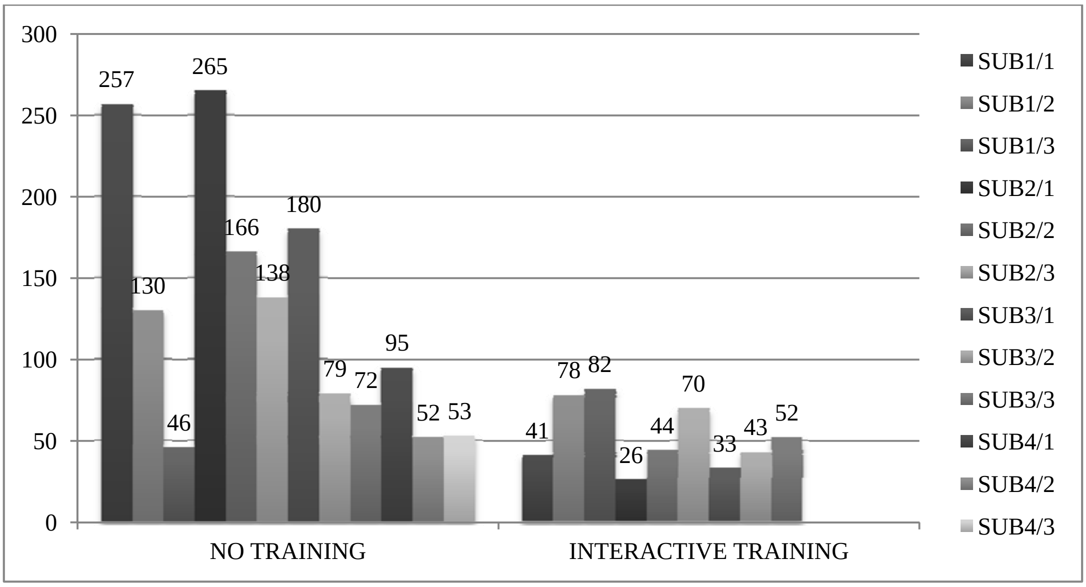

Fig. 3. MLT Completion Comparison.

3. Discussion

The results featured in fig. 1 emphasize the effectiveness of the interactive training paradigm for the vOICe. In addition, figs. 2 and 3 present a surprising result of decreasing effectiveness across trials for interactive training. These results prompt two questions:

(1) Why did the interactive training result in faster completion times?

(2) Why did completion times slow in subsequent MLT trials for the interactive trained group, particularly in contrast to the subjects who were not trained?

There are two related ways of answering these questions particularly relevant to psychology: (i) the possible activation of "social" cognition in the interactive training, and (ii) that perceptual learning strategies are task-specific across the different training techniques. In addition, there is a related philosophical explanation for these results; our minds integrate with our environment in such a way that when the environment responds to our activities we are able to synchronize with it, thereby activating more channels of mental activity.

3.1. *Social Cognition*

The utilization of social cognition in interactive training may help answer question (1) (Stich and Nichols 1992). For it has shown that subjects interactively trained with other subjects are engaging in a more natural and effective learning paradigm (Hobson 2002). This is likely because games that involve interaction with intentional subjects acquire a larger set of skills (Merabet and Sanchez 2009). If so, this is further evidence that those using SSDs are able to distinguish between interactions with intentional subjects and mere objects (Auvray and Myin 2009). This explanation suggests that SSD use allows for "social" cognition, and that the interactive training paradigm of SSD tag may exploit the suggested benefits of better skill acquisition that comes along with it.

The second set of results of slower completion times across trials for the interaction training set might also be explained by the utilization of social cognition. For subjects may recognize that in the MLT they are interacting with objects rather than the intentional subjects of the interactive tag training. Hence, it may be that perception modulates across two types: perceiving passive object types and interactive subject types.[3] This suggests, as will be discussed below, that perceptual training on SSDs is task-sensitive.

3.2. *Perceptual Training*

Perceptual training studies are also relevant in response to question (1). They indicate that the ratio of specificity to transfer (i.e., how much one's training transfers to a different task versus how specific the training remains to the training task itself) is highly affected by task relevance (Jeter et al. 2010). In the case of our study, the passive training method likely down weighted the task-irrelevant channels (such as active bodily engagement with the sounds) as subjects played tic-tac-toe, making it difficult to transfer the learning achieved through passive training to the more active MLT. For instance, bodily movements and spatial maps were not relevant to the passive task, and were potentially pruned as irrelevant to the SSD's auditory signals. Interactive training, on the other hand, provided the broadest possible range of relevant stimuli among the conditions we used. Thus, a broader range of perceptual skills was allowed to transfer to the MLT. Subjects learned to associate movement, spatial maps, and environmental interaction as highly related to the signals coming from the vOICe.

The ratio of specificity to transfer may also help explain question (2) about why subjects in the no training group showed consistent improvement across trials, whereas passive and actively trained subjects showed no improvement. No training subjects were pruning task-irrelevant channels, such as passive utilization of the sounds, to the active MLT task, whereas passive and active trainees had already optimized the perceptual learning channels (i.e., channels used to interpret signals given by the SSD) and did not show further perceptual learning after the three MLT trials. Interactively trained subjects, moving from a rich environment (broad range of relevant stimuli) to a less rich environment in the MLT, may have begun to prune channels used in their training but absent in the new task. This pruning may have caused the small but significant slowing in MLT task time as channels were pruned across the three trials.

3.3. Extended Mind Thesis

The extended mind view suggests that the best training environment is one that strikes back—that recursively affects a subject's actions and thoughts. SSD tag is just such a kind of training paradigm: users who seek to tag another subject recognize and eventually experience the fact that they too are a target to be tagged. The extended mind view suggests that the more opportunity for a subject to actualize their extended activity in their environment, the more adept at learning their activity they will be. "[T]he human organism," writes Clark, "is linked with an external entity in a two-way interaction, creating a coupled system that can be seen as a cognitive system in its own right" (2008, 222). The extended mind view then might explain adept SSD use after SSD tag training by appealing to the notion that subjects, their devices, and their environment become embedded in a single mental activity, creating a more hospitable learning environment than mere active or passive use of the device.

The extended mind thesis might also suggest an answer to question (2). The increasingly poor results only shown by interactive trained subjects suggest that the MLT environment was increasingly inhospitable to the kind of SSD function that subjects were trained to employ. This explanation emphasizes the effectiveness of interactive training. For though the results from these subjects were by far the quickest, the MLT environment was not an ideal environment for the utilization of their training. In sum, the interactive training was so effective that even in inhospitable environments it outperformed its competitors.

4. Conclusion

The active training, which is most closely associated with the enactivist account of perception, specifically prepared subjects for the MLT. However, active trained subjects performed as poorly as the passive trained and naïve. By contrast, subjects interactively trained performed better even though the MLT environment was inhospitable to this training paradigm. The extended mind theory most closely associated with this training is thus strongly suggested by these results.

Notes

1. For an exception see Merabet and Sanchez, 2009.
2. The MLT completion times of one subject had to be eliminated because of significant groping for the luminescent discs rather than using the sounds emitted from the device to locate them.
3. This distinction is made by Auvray and Myin (2009).

Works Cited

Auvray, Malika, and Erik Myin. “Perception with Compensatory Devices: From Sensory Substitution to Sensorimotor Extention.” *Cognitive Science* 33.6 (2009): 1036-58.

Auvray, Malika, Sylvain Hanneton, and Kevin O’Regan. “Learning to Perceive with a Visuo-Auditory Substitution System: Localization and Object Recognition with the Voice.” *Perception* 36.3 (2007): 416-30.

Chebat, Daniel R., Constant Rainville, Ron Kupers, and Maurice Ptito. “Tactile-‘Visual’ Acuity of the Tongue in Early Blind Individuals.” *Neuroreport* 18.18 (2007): 1901-04.

Clark, Andy. *Supersizing the Mind*. New York: Oxford UP, 2008.

Hobson, Peter. *The Cradle of Thought: Challenging the Origins of Thinking*. New York: Oxford UP, 2002.

Jeter, Pamela E., Barbara Anne Dosher, Liu Shiau-Hua, and Lu Zhong-Lin. “Specificity of Perceptual Learning Increases with Increased Training.” *Vision Research* 50 (2010): 1928-40.

Meijer, Peter B. L. “An Experimental System for Auditory Image Representations.” *IEEE Transactions Biomedical Engineering* 39.2 (1992): 112-21.

Merabet, Lotfi B., and Jaime Sanchez. “Audio-Based Navigation Using Virtual Environments: Combining Technology and Neuroscience.” *AER Journal: Research and Practice in Visual Impairment and Blindness* 2.3 (2009): 128-37.

Nöe, Alva. *Action in Perception*. Cambridge: MIT, 2006.

Proulx, Michael J., Petra Stoerig, Eva Ludowig, and Inna Knoll. “Seeing ‘Where’ through the Ears: Effects of Learning-by-Doing and Long-Term Sensory Deprivation on Localization Based on Image-to-Sound Substitution.” *PLoS ONE* 3.3 (2008): e1840.

Ptito, Maurice, Solvej M. Moesgaard, Albert Gjedde, and Ron Kupers. “Cross-Modal Plasticity Revealed by Electrotactile Stimulation of the Tongue in the Congenitally Blind.” *Brain* 128.3 (2005): 606-14.

Stich, Stephen, and Shaun Nichols. “Folk Psychology: Simulation or Tacit Theory.” *Mind & Language* 7.1 (1992): 35-71.