Relativism, Fascism, and the Question of Ethics in Constructivism
by Ernst von Glasersfeld

Heinz von Foerster Congress – Part 2

Reflexivity and Eigenform
by Louis Kauffman

Conversation by Klaus Krippendorff

Conversation, Individuals and Concepts by Bernard Scott

Consciousness as Symbolic Construction by Hans-Herbert Kögler

Book review of Radical Enactivism and Disorders of Volition

Publication Review
Description

Constructivist Foundations (CF) is an independent academic peer-reviewed e-journal without commercial interests. Its aim is to promote scientific foundations and applications of constructivist sciences, to weed out pseudoscientific claims and to base constructivist sciences on sound scientific foundations, which do not equal the scientific method with objectivist claims. The journal is concerned with the interdisciplinary study of all forms of constructivist sciences, such as radical constructivism, biology of cognition, cybersemiotics, enactive cognitive science, epistemic structuring of experience, non-dualizing philosophy, second order cybernetics, and theory of autopoietic systems.

Aim and Scope

The basic motivation behind the journal is to make peer-reviewed constructivist papers available to the academic audience free of charge. The constructive character of the journal refers to the fact that the journal publishes actual work in constructivist sciences rather than work that argues for the importance or need for constructivism. The journal is open to (provocative) new ideas that fall within the scope of constructivist approaches and encourages critical academic submissions to help sharpen the position of constructivist sciences. The common denominator of constructivist approaches can be summarized as follows.

According to constructivist approaches, it is futile to claim that knowledge approaches reality; reality is brought forth by the subject rather than passively received.

Constructivist approaches entertain an agnostic relationship with reality, which is considered beyond our cognitive horizon; any reference to it should be refrained from.

Therefore, the focus of research moves from the world that consists of matter to the world that consists of what matters.

Constructivist approaches focus on self-referential and organizationally closed systems; such systems strive for control over their inputs rather than their outputs.

With regard to scientific explanations, constructivist approaches favor a process-oriented approach rather than a substance-based perspective, e.g., living systems are defined by processes whereby they constitute and maintain their own organization.

Constructivist approaches question the Cartesian separation between objective world and subjective experience.

Constructivist approaches emphasize the “individual as personal scientist” approach; sociality is defined as accommodating within the framework of social interaction.

Finally, constructivist approaches ask for an open and less dogmatic approach to science in order to generate the flexibility that is needed to cope with today’s scientific frontier.

Publisher

Alexander Riegler (editor-in-chief) · Postal address Center Leo Apostel for Interdisciplinary Research, Vrije Universiteit Brussel, Krijgskundestraat 33, B-1160 Brussels, Belgium

Physical web server University of Vienna, A-1010 Vienna, Austria · ISSN 1782-348X

Website http://www.univie.ac.at/constructivism/journal

Acknowledgment to Reviewers of Volume 4

We thank the following external reviewers for taking the time to review manuscripts submitted to Constructivist Foundations Volume 4: Mariaelena Bartesaghi, Lea Sugman Bohinc, Gary M. Boyd, Andrew Brook, Richard Buttny, Martin Coleman, Henk de Regt, Igor Douven, Georgianna Gardner, Hugh Gash, Dagmar Hoffmann, Alfred Inselberg, Stephen G. Loftis, Solomon Marcus, Winfried Nöth, Konrad Ott, Paul Pangaro, Ana Pasztor, Andreas Quale, Armin Scholl, Leslie P. Steffe, Peter Strasser, Jean Paul van Bendegem, Stefan Weber.

Cover Art


Submissions are continuously received

Send all material to Alexander Riegler <ariegler@vub.ac.be>
Relativism, Fascism, and the Question of Ethics in Constructivism

Ernst von Glasersfeld ◦ University of Massachusetts <evonglas@hughes.net>

Purpose – Radical constructivism holds that experiential reality is created by each individual. As a way of thinking, it unquestionably belongs to the theories of knowledge that are called “subjectivist” and “relativist”. This paper deals with the Italian philosopher Adriano Tilgher’s analysis of the relation between relativism and fascism and examines the possible impact of this connection on constructivism and its view of ethics.

Approach – Conceptual analysis and the demonstration of a contradiction in Tilgher’s argumentation. Findings – A review of the ethics inherent in Kant’s categorical imperative shows a tendency towards a subjectivism that may hint at anarchism but in no way implies fascism. Implications – This investigation should make it more difficult for critics of constructivism to relate it to fascism. Key words – Relativism, subjectivism, ethics, categorical imperative.

Introduction

At the beginning of Probabilitismo, the first comprehensive statement of his philosophy, the mathematician Bruno de Finetti quoted a passage that intrigued me no end: “Truth no longer resides in the imaginary equality of an idea and what lies outside the mind; if it is outside the mind there is no way to touch and apprehend it. Truth arises in the act of thinking of a thinker. […] Thought is not a mirror that reflects unaltered a reality external to us; thought is simply a biological function, a means to orient oneself in life” (de Finetti 1989: 169).

The quotation, given its date, could not but surprise a constructivist. It was actually written ten years earlier by Adriano Tilgher in his essay, Relativisti contemporanei (“contemporary relativists”), published for the first time in 1921. The essay is a critical review of the epistemological ideas of Hans Vaihinger, Albert Einstein, Louis Rougier (an interpreter of Henri Poincaré), and Oswald Spengler – authors that have interested me a great deal because they expressed ideas that anticipated constructivism. Through a friend in Italy, I managed to get a tattered copy of the fourth edition of Tilgher’s essay (Tilgher 1923).

Relativism and politics

Adriano Tilgher (1887–1941) was a prolific thinker. He published more than a dozen books on subjects as varied as the Greek philosophers, pragmatism, knowledge and reality, esthetics, and the theatre. The last two sections of the essay on “contemporary relativists” are entitled Today’s Idealism and Relativism and Revolution. There he sets out his ideas concerning the connection of subjectivism with politics and ethics.

“The success of today’s idealism, which is but sublimated and essentialized Kant, springs from the fact that its focus on individual action satisfies, more than any other philosophy, the generation that has been involved in the war” (Tilgher 1923: 76).

“Individuals can no longer act in the hope that their actions […] will crystallize into institutions. […] Their only reason to act can be the wish to escape by means of acting from the desperation of a cursed world, a world that no longer permits even the flight into the desert. It is action for action, action for the sake of action, ad infinitum, … and this action can have only a negative purpose: the overthrow of existing things” (Tilgher 1923: 90–91).

As a result of the war, he explains, the authority of the state gave way to the “proletarian assault” (Tilgher 1923: 76) that led to a fascist movement that proclaimed that the political movement can be impeded by rational arguments. One of the reviewers of my article criticized it for not providing arguments against a “Führer’s” subjective authoritarianism. As I try to show, there are no rational arguments for or against subjective ethical precepts (cf. initial quotation from Reich).

“A systemic/constructivist pedagogy […] accepts the constructivity of world views, but cannot decide definitely to accept one that would rigidly limit its possibilities of observation. Although constructivists, too, have to make everyday and political decisions, the observer’s freedom means so much to them that they prefer to suffer mutually incompatible opinions than submit to the tyranny of a single view that is to serve as construct for all.” Kersten Reich (1996: 21).
Having found a plausible source of subjectivism in Kant, Tilgher proceeds to argue that it is only a small step from there to the justification of individual action and to anarchism and fascism. The basis for subjectivism in Kant is indeed solid. I shall argue, however, that Kant’s view may be compatible with anarchism but certainly not with fascism. Statements such as the following leave no doubt about Kant’s position in that regard: “Der Mensch denkt mit seinem Verstand ursprünglich, und er schafft sich also seine Welt” [With their understanding (mind, reason) human beings generate original thoughts and thus create their world] (Kant 1798: 71).

No doubt this “creating” points to an activism of sorts, but what Kant is talking about are mental operations, not the physical actions that were at the core of the fascist doctrine. Kant’s subjectivism also emerges from his theory of ethics to which I will return in the next section.

According to Tilgher, in idealism, which he considers to be the root of relativism, an act of thought at every moment constitutes reality and thus generates everything that can exist (Tilgher 1923: 80). On the basis of this transcendence of experiential reality, he argues that there can be no “progress” because this would imply a reality more complete than the one that seems real at the moment.

“How could one find a difference of value between the various products of the mind, if all of them are complete and all that there is at a given moment? And how, without differences of value, would it be possible to develop notions of progress or history? The pure act of the mind thus breaks up into countless centers of action that are absolutely equivalent. This absolute activism leads to absolute relativism, individualism, and anarchism” (Tilgher 1923: 80–81).

Relativism cannot lead to both anarchism and fascism at the same time. Anarchism is incompatible with Mussolini’s doctrine, as well as with Hitler’s. Both require the inflexible structure of a powerful state that can enforce an ethic and exercise rigid control over individuals’ leeway. The Concise Oxford Dictionary of Politics (McLean 1996) defines anarchism as “the view that society can and should be organized without a coercive state.” This is the aspect Paul Feyerabend had in mind when he described himself as an “epistemological anarchist” (Feyerabend 1978: 21, note 12). He explains that he is not in favor of the popular, belligerent form of anarchism but rather of the benign form that simply objects to the imposition of absolute rules. This is in tune with Theodor Adorno who, in his famous essay Erziehung nach Auschwitz (“Education after Auschwitz,” 1967) wrote: “The only veritable power against the principle of Auschwitz would be autonomy, the power of reflection, of self-determination, of not-going-along.” And that is how I intend anarchism. The fascist state was diametrically opposed to such individualism. It considered itself all-powerful and unquestionable and had no qualms about coercion. “Mussolini ha sempre ragione” (M. is always right) was one of the slogans in Italy; and because there was a humorous opposition until Hitler took over, people used to say: “Anche se sbaglia!” (also when he’s wrong). That is why countless anarchists were persecuted and kept in prison by Hitler and Mussolini.

**Ethics and constructivism**

The way I come to the conclusion that there is a subjectivist/anarchist element in Kant’s theory of ethics springs from his formulation of the “Categorical Imperative,” which crops up at various places in his works. The wording is not always quite the same but the variations are all compatible with the basic statement in Kant’s Grundlegung zur Metaphysik der Sitten (“Foundations of the metaphysics of morals”):

“Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, daß sie ein allgemeines Gesetz werde” (Kant 1785: 421).

“Act only according to that maxim by which you can at the same time will that it should become a universal law” (Translation: Encyclopedia Britannica 2004).

This is the best English translation I have seen; except that “allgemein” would be better translated as “general.” Many authors abbreviate the Imperative so that it comes to imply not a desired maxim of action but an already existing absolute law. My Italian philosophical dictionary (Abbagnano 1964: 461) goes so far as to say that the Categorical Imperative is the command for an action that “is good in itself.” This is not what Kant intended. He explicitly said that the maxim should be such that you could will that it should become a general law. The maxim you wish others would adopt is yours, it is your subjective intuition. In other words, you should act only in ways in which, according to you, other people, too, should act. Kant was not one to use words in a slovenly fashion. When he wrote “you can will” he intended a wish, not the statement of an accepted fact; and as a wish it entails a fundamental subjectivity. In his Canon of Pure Reason (1787: 833), Kant asks the rhetorical question: “What ought I to do?” He explains that it is usually a practical question and as such may be answered in the domain of reason; but it can also be a moral question, that is, a question with which reason cannot deal.

Kant, of course, also wants to show how general ethical rules can be obtained: but, as he indicated by the title of his “Foundations of the metaphysics of morals,” to do this he has to turn to metaphysics. He leans on the notion of God and of an afterlife, and thus turns to traditional, a-rational metaphysics, which the theory of knowledge laid out in his Critique of pure reason so successfully avoided.

Heinz von Foerster’s precept that you should always say “I should!” and never “You should!” (Foerster 1993: 354) similarly implies that ethical prescriptions must be personal and must not involve coercion. I fully agree with this.

A woman who was a teenager in the 1960s, the hippy era, explained to me that there was really no problem about developing your own ethics: you mistrusted any and all of the rules your parents tried to impose, but you learned to avoid what had unpleasant consequences. I was stunned. It was an unforeseen application of the notion of first-level viability, namely that on the level of action as well as that of cognition anything that works is retained as workable. It is the basis for the generation of our subjective rules of action. It may be extended to cover the suggestion I have frequently made (Glaserfeld 1983: 223, 1995: 120), namely that we use our construction of “others,” who seem to have action rules similar to our own, as the foundation of a second-order viability. But it does not lead to an extra-personal ethics; just as no amount of studying works of art can provide an objective scale in esthetics.
Conclusion

The accusation that relativism or, indeed, constructivism prepares the way to fascism is at best a thoughtless one. If general ideas of ethics could determine the choice of political systems, the Christian version of ethics that was dominant in Italy and Germany should have prevented fascism and Nazism. But it was not only Pacelli, the pope of the Catholic church, but thousands of Catholic priests and protestant ministers, too, who failed to raise their voice against the totalitarian movement.

Social/political movements tend to develop their own ethics and, if they emphasize precepts such as “my country right or wrong” or turn a party or its leader into an idol that demands unquestioning obedience, the way is open to collective actions without ethical considerations of any sort. Obedience is then the only “good” and it is claimed to relieve the individual of responsibility.

Rational models of epistemology can determine neither ethical nor esthetical standards, which are part of metaphysics, a domain that, as Wittgenstein (1933: 6.421) clearly stated, is inaccessible to reason. Radical constructivism may, however, imply a starting-point for the development of an ethical system, not the system itself. RC suggests that if actions of “others” that you have constructed – a construction that is not ad lib but is constrained by facts of experience – can be interpreted as similar to your own, they may confirm a second-order viability. But this can be concluded only if the others act without coercion. This seems to fit well with the principle of anarchism, but not at all with the authoritarian dictates of fascism.

Acknowledgment

I thank Marie Larochelle, Marco Bettoni, and Alexander Riegler for very helpful comments on a draft of this paper, and Siegfried Schmidt for help with references.

Notes

1. All translations are mine unless stated otherwise.
2. Critics of radical constructivism have suggested that this “use” of others contra-venes Kant’s Categorical Imperative (Ott 1995: 394). This is a misguided notion. In the first place, what they have in mind is Kant’s statement: “Act so as to treat humanity, whether in your own person or in another, always as an end, and never as only a means” (Kant 1785: 429), and this is not the “categorical” imperative; Kant formulated it as a “practical” imperative. Second, and more important, is the fact that using others in order to raise the level of viability of your own thinking and acting is not using them as “things,” it merely uses your interpretation of their doings.

References

Kant I. (1798) Der Streit der Fakultäten. Akademie-Ausgabe, Volume VII.


Received: 27 March 2009
Accepted: 2 July 2009
The Shape of Process

1. Introduction

“Reflexive” is a term that refers to the presence of a relationship between an entity and itself. One can be aware of one’s own thoughts. An organism produces itself through its own action and its own productions. A market or a system of finance is composed of actions and individuals, and the actions of those individuals influence the market just as the global information from the market influences the actions of the individuals. Here it is the self-relations of the market through its own structure and the structure of its individuals that moves its evolution forward. Nowhere is there a way to cut an individual participant from the market effectively and make him into an objective observer. His action in the market is concomitant to his being reflexively linked with that market. It is just so for theorists of the market, for their theories, if communicated, become part of the action and decision-making of the market. Social systems partake of this same reflexivity, and so does apparently objective science and mathematics. In order to see the reflexivity of the practice of physical science or mathematics, one must leave the idea of an objective domain of investigation in brackets and see the enterprise as a wide-raying conversation among a group of investigators. Then, at once, the process is seen to be a reflexive interaction among the members of this group. Mathematical results, like all technical inventions, have a certain stability over time that gives them an air of permanence, but the process that produces these novelities is every bit as fraught with circularity and mutual influence as any other conversation or social interaction.

How then, shall we describe a reflexive domain? It is the purpose of this paper to give a very abstract definition that nevertheless captures what I believe to be the main conceptual feature of reflexivity. We then immediately prove that eigenforms, fixed points of transformations, are present for all transformations of the reflexive domain. This will encourage us and will give us pause to think further about the relationship of reflexivity and eigenform.

The existence of eigenforms will encourage us, for we have previously studied them with the notion that “objects are tokens for eigenbehaviors.” Eigenforms are the natural emergence of those tokens by way of recursion. So to find the eigenforms dictated by a larger concept is pleasing. The existence of fixed points for arbitrary transformations shows us that the domain we have postulated is indeed very wide. It is not an objectively existing domain. It is a clearing in which structures can arise and new structures can arise. A reflexive domain is not an already-existing structure. To be what it claims to be, a reflexive domain must be a combination of an existing structure and an invitation to create new structures and new concepts. The new will become platforms from which further flights of creativity can be made. Thus in the course of examining the concept of reflexivity we will find that the essence of the matter is an opening into creativity; and that will become the actual theme of this paper.
This essay begins with a discussion of the notion of “eigenform” as pioneered by Heinz von Foerster in his papers (Foerster 1981a–c) and explored in papers by the author (Kauffman 1987, 2003, 2005). We include some of the material from (Kauffman 2005) in this paper for the sake of completeness. In (Foerster 1981a) the familiar objects of our existence can be seen to be nothing more than tokens for the behaviors of the organism, creating apparently stable forms.

In this view, the object is both an element of a world and a token or symbol for the process of its production/observation.

An object, in itself, is a symbolic entity, participating in a network of interactions, taking on its apparent solidity and stability from these interactions. We ourselves are such objects: we, as human beings, are “signs for ourselves,” a concept originally developed by the American philosopher, Charles S. Peirce (Kauffman 2001). Eigenforms are mathematical companions to Peirce’s work.

In an observing system, what is observed is not distinct from the system itself, nor can one make a complete separation between the observer and the observed. The observer and the observed stand together in a coalescence of perception. From the stance of the observing system, all objects are non-local, depending upon the presence of the system as a whole. It is within that paradigm that these models begin to live, act and enter into conversation with us.

After this journey into objects and eigenforms, we take a wider stance and consider the structure of spaces and domains that partake of the reflexivity of object and process. In Section 6 we give a definition of a reflexive domain. Our definition populates a space (domain) with entities that could be construed as objects, and we assume that each object acts as a transformation on the space. Essentially this means that given entities $A$ and $B$, there is a new entity $C$ that is the result of $A$ and $B$ acting together in the order $AB$ (so that one can say that “$A$ acts on $B$” for $AB$ and “$B$ acts on $A$” for $BA$). This means that the reflexive space is endowed with a non-commutative and non-associative algebraic structure. The reflexive space is expandable in the sense that whenever we define a process, using entities that have already been constructed or defined, then that process can take a name, becoming a new entity/transformation of a space that is expanded to include itself. Reflexive spaces are open to evolution over time as new processes are invented and new forms emerge from their interaction.

Remarkably, reflexive spaces always have eigenforms for every element/transformation/entity in the space! The proof is simple but requires discussion.

Given $F$ in a reflexive domain. Define $G$ by $Gx = F(xx)$.

Then $GG = F(GG)$ and so $GG$ is an eigenform for $F$.

Just as promised, in a reflexive domain, every entity has an eigenform. From this standpoint, one should start with the concept of reflexivity and see that from it emerge eigenforms. Are we satisfied with this approach? We are not. In order to start with reflexivity, we need to posit objects and processes. As we have already argued in this essay, objects are tokens for eigenbehaviors. And a correct or natural beginning is a process where objects are seen as tokens of processes.

By now the reader begins to see that the story we have to tell is a circular one. We give a way to understand this circularity in our last section, where we discuss creativity in recursive processes and the emergence of novelty.

The paper continues in Section 6 by studying an allied concept that we call a magma. A magma is a domain with a binary operation $*$ that allows one to combine elements $a$ and $b$ of the domain to form a new element $a * b$ of that domain. In the magma each element $a$ is also a mapping of the domain to itself via left combination: $x \mapsto a * x$. We assume that each such transformation preserves the structure of the combinatory operation. Magmas are very close in concept to reflexive domains. We define the notion of a reflexive magma and show that such magmas satisfy a fixed point theorem and so contain eigenforms. In Section 7 we show how magmas arise naturally in the context of knot theory and a theory of knot sets. Sections 8 and 9 discuss the relationships of reflexivity with the lambda calculus of Church and Curry and with Cantor’s diagonal argument and the Russell paradox. Section 10 is a minimalist discussion of self-reference and reflexivity in relation to the conceptualization of a universe that comes to observe itself. Section 11 is an account of a computer experiment with a variant of the Life cellular automaton of John H. Conway. In this variant that we have discovered, Life, the recursions lead to self-sustaining processes with very long evolutionary patterns. We show how examples of novel phenomena arise over the course of large time scales. This example will be a later springboard for the discussion of the emergence of novelty from deterministic processes. Here, it is an example showing how the course of a process is just as important as its eigenform or infinite concatenation.

The paper ends with a discussion of the wider context of reflexivity. We are acutely aware that this paper about reflexivity only gives certain conceptual tools and does not yet address the actuality of the reflexive condition of persons and observers who are inextricably part of the universes that they hope to study. In so doing they will adopt points of view and these very points of view will create patterns, new forms, objects of study and will act as a veil over the original intent. It is only through working with many points of view and many investigations that the particularities of single lenses will begin to fall away and a wider understanding will emerge.

2. Objects as tokens for eigenbehaviors

In his paper Objects as Tokens for Eigenbehaviors, von Foerster (1981a) suggests that we think seriously about the mathematical structure behind the constructivist doctrine that perceived worlds are worlds created by the observer. At first glance such a statement appears to be nothing more than solipsism. At second glance, the statement appears to be a tautology, for who else can create the rich subjectivity of the immediate impression of the senses? At third glance, something more is needed. In that paper he suggests that the familiar objects of our experience are the fixed points of operators. These operators are the structure of our perception. To the extent that the operators are shared, there is no solipsism in this point of view. It is the beginning of a mathematics of second order cybernetics.

Consider the relationship between an observer $O$ and an “object” $A$. The key point about the observer and the object is that “the object remains in constant form with respect
to the observer.” This constancy of form does not preclude motion or change of shape. Form is more malleable than the geometry of Euclid. In fact, ultimately, the form of an object is the form of the distinction that “it” makes in the space of our perception. In any attempt to speak absolutely about the nature of form we take the form of distinction for the form (paraphrasing Spencer-Brown 1969). It is the form of distinction that remains constant and produces an apparent object for the observer. How can you write an equation for this? The simplest route is to write

\[ O(A) = A. \]

The object \( A \) is a fixed point for the observer \( O \). The object is an eigenform. We must emphasize that this is the most schematically possible description of the condition of the observer in relation to an object \( A \). We only record that the observer as an actor (operator) manages through his acting to leave the (form of) the object unchanged. This can be a recognition of the symmetry of the object but it also can be a description of how the observer, searching for an object, makes that object up (like a good fairy tale) from the very ingredients that are the observer herself. This is the situation that Heinz von Foerster has been most interested in studying. As he puts it, if you give a person an undecidable problem, then the answer that he gives you is a description of himself. And so, by working on hard and undecidable problems we go deeply into the discovery of who we really are. All this is symbolized in the little equation \( O(A) = A. \)

And what about this matter of the object as a token for eigenbehavior? This is the crucial step. We forget about the object and focus on the observer. We attempt to “solve” the equation \( O(A) = A \) with \( A \) as the unknown. Not only do we admit that the “inner” structure of the object is unknown, we adhere to whatever knowledge we have of the observer and attempt to find what such an observer could observe based upon that structure.

We can start anew from the dictum that the perceive and the perceived arise together in the condition of observation. This is a stance that insists on mutuality (neither perceive nor the perceived causes the other). A distinction has emerged and with it a world with an observer and an observed. The distinction is itself an eigenform.

3. Compresence and coalescence

We identify the world in terms of how we shape it. We shape the world in response to how it changes us. We change the world and the world changes us. Objects arise as tokens of a behavior that leads to seemingly unchanging forms. Forms are seen to be unchanging through their invariance under our attempts to change, to shape them.

For an observer there are two primary modes of perception — compresence and coalescence. Compresence connotes the coexistence of separate entities together in one including space. Coalescence connotes the one space holding, in perception, the observer and the observed, inseparable in an unbroken wholeness. Coalescence is the constant condition of our awareness. Coalescence is the world taken in simplicity; compresence is the world taken in apparent multiplicity.

This distinction between compresence and coalescence, drawn by Henri Bortoft (1971), can act as a compass in traversing the domains of object and reference. Eigenform is a first step towards a mathematical description of coalescence. In the world of eigenform, the observer and the observed are one in a process that recursively gives rise to each.

4. The eigenform model

We have seen how the concept of an object has evolved to make what we call objects (and the objective world), processes that are interdependent with the actions of observers. The notion of a fixed object has become a notion of a process that produces the apparent stability of the object. This process can be simplified in a model to become a recursive process where a rule or rules are applied time and time again. The resulting object of such a process is the eigenform of the process, and the process itself is the eigenbehavior.

In this way we have a model for thinking about object as token for eigenbehavior. This model examines the result of a simple recursive process carried to its limit. For example, suppose that

\[ F(X) = \]

That is, each step in the process encloses the results of the previous step within a box. Here is an illustration of the first few steps of the process applied to an empty box, \( X \):

\[ \]

If we continue this process, then successive nests of boxes resemble one another, and in the limit of infinitely many boxes, we find that

\[ X = F(F(\ldots)) = \]

The infinite nest of boxes is invariant under the addition of one more surrounding box. Hence this infinite nest of boxes is a fixed point for the recursion. In other words, if \( X \) denotes the infinite nest of boxes, then

\[ X = F(X). \]

This equation is a description of a state of affairs. The form of an infinite nest of boxes is invariant under the operation of adding one more surrounding box. The infinite nest of boxes is one of the simplest eigenforms.

Remark. On reading the above description of the limiting process

\[ X \rightarrow F(X) \rightarrow F(F(X)) \rightarrow \ldots \]

the reader may find herself thinking along the following lines: “Doesn’t he mean to put those three dots in the nested boxes on the outside of the boxes rather than on the inside? After all,
the operation $F$ surrounds $X$ with a square, so at each stage, a square is added from the outside. Shouldn’t the picture then be like this one below?

```
  ...
  ...
  ...
```

“I have illustrated the new picture with the three dots on the left, the right, the top and the bottom to show how in this way of thinking the nest of boxes grows outward and consequently it grows in all these directions.

If we take this construction to infinity, then it will either fill the plane with boxes, or the widths between successive boxes will have to grow smaller and smaller, just as, with the three dots inside, you had to make the boxes smaller and smaller. But really, this second picture is quite different from the first picture. In fact if we do make the second picture, and imagine that it is a solution to the equation $F(X) = X$, it does not seem to be a solution! Look at the picture below.”

```
  ...
  ...
  ...
```

“Now I have put a box around the outwardly growing infinite nest of boxes, but this means that I have allowed an infinite number of boxes to grow there (going out but staying in a finite amount of space by crowding one next to another) and then I put one more box around all of them. The result is not the same! This is a new form of boxes.

So with the outward growth, I make new infinites, but I do not solve the equation $X = F(X)$. Now I see what you were doing with the inward nest of boxes. You let it grow inwardly and obtained a limit form that did not see the one box more that you put around the outside. I had to try this other method in order to see what you were doing. And I am sure that other readers will have to experiment in this way and in new ways to really understand this construction of eigenforms.”

Comment on the remark. Indeed the patient reader was right that there is more than one way to go to infinity. A simpler example can be seen in the equation $x = ax$ where we solve it by letting $x = aaa...$, an infinite repetition of $a$’s going off to the right.

Then $ax = a(aaa...) = aaaa... = a... = x$.

But if we do it in the other order and take $xa$, we find that

$$xa = aaaa...a$$

which means an infinite row of $a$’s followed by one more $a$.

And we see that in this way of thinking $xa$ is not equal to $ax$.

Similarly, $y = ...a$$aa$ is not a solution to $ay = y$ but it is a solution to $yu = a$. This may seem a bit strange and abstract, so it is better to think with the boxes (I think). But in ordinary mathematics we use this same sort of infinite construction. For example, we write

$$x = 1 + a + aa + aaa + aaaa + ...$$

and rewrite it in the form

$$x = 1 + a(1 + a + aa + aaa + ...) = 1 + ax$$

and conclude that

$$x - ax = 1,$$

$$x(1 - a) = 1,$$

$$x = 1 / (1 - a).$$

Hence

$$1 / (1 - a) = 1 + a + aa + aaa + aaaa + ...$$

Here we are using $ax$ for the product of a with itself, so these can be numbers. And one can verify that indeed if $a$ is a number and the absolute value of $a$ is less than one, then this formula is true. For example, if $a = 1 / 2$, then

$$1 / (1 - 1 / 2) = 2$$

and the formula asserts that

$$2 = 1 + 1/2 + 1/4 + 1/8 + 1/16 + ...$$

This is true, and the reader should ask herself how she knows that it is true! The reader will also be interested in seeing what happens when $a$ is bigger than or equal to 1 in absolute value. For example, if $a = 2$, then our formula would seem to say that

$$-1 = 1 + 2 + 4 + 8 + 16 + 32 + ...$$

Is there some truth in this absurdity?

We can see what has actually happened by making a closer analysis.

Let $X = 1 + 2 + 4 + ... + 2N$ where $2N$ means 2 multiplied by itself $N$ times. Then we have

$$X = 1 + 2(1 + 2 + ... + 2^{N-1}),$$

$$X = 1 + 2\times2^{N+1} - 2(2^N),$$

So we have

$$X - 2X = 1 - 2^{N+1},$$

which is the same as saying

$$X = -1 + 2^{N+1}.$$

Do you see what has happened? We are interested in finding out what happens when $N$ goes to infinity. But here if we ignore the term $2^{N+1}$ we will get the wildly wrong answer of $-1$.

You have to take infinity with a grain of salt as well as looking at it as the vastness of all the grains of sand on the beach. End of comment.

A further comment. Perhaps you thought that we showed that the equation: $-1 = 1 + 2 + 4 + ...$ is wrong. There is a point of view in which it is right! Consider that in binary arithmetic we represent 1 by 1, 2 by 10, 4 by 100, 8 by 1000 and so on. Then $1 + 2 + 4 + 8$ is represented in binary by 1111, and when you add 1 to 1111 you find a series of carrys taking you to the answer 10000. Suppose you had a computer that could only handle binary numbers up to four bits. Then when you added 1 to 1111 you would get 0000, since the computer would throw away the last bit. In this sense 1111 represents $-1$ in such a limited computer, and in the same way the infinite sum $1 + 2 + 4 + 8 + ...$ represents $-1$ in an infinite computer that is not prepared to have bits beyond the first infinity!

In the process of observation, we interact with ourselves and with the world to produce stabilities that become the objects of our perception. These objects, like the infinite nest of boxes, may go beyond the specific properties...
of the world in which we operate. They attain their stability through the limiting process that goes outside the immediate world of individual actions. We make an imaginative leap to complete such objects to become tokens for eigenbehaviors. It is impossible to make an infinite nest of boxes. We do not make it. We imagine it. And in imagining that infinite nest of boxes, we arrive at the eigenform.

The leap of imagination to the infinite eigenform is a model of the human ability to create signs and symbols. In the case of the eigenform \( X = F(X) \), \( X \) can be regarded as the name of the process itself or as the name of the limiting process. Note that if you are told that

\[
X = F(X),
\]

then, substituting \( F(X) \) for \( X \), you can write

\[
X = F(F(X)).
\]

Substituting again and again, you have

\[
X = F(F(F(X))) = F(F(F(F(F(X))))) = \ldots
\]

The process arises from the symbolic expression of its eigenform. In this view, the eigenform is an implicate order for the process that generates it. (Here we refer to implicate order in the sense of David Bohm 1980).

Sometimes one stylizes the structure by indicating where the eigenform \( X \) reenters its own indicational space with an arrow or other graphical device. See the picture below for the case of the nested boxes.

Does the infinite nest of boxes exist? Certainly it does not exist on this page or anywhere in the physical world with which we are familiar. The infinite nest of boxes exists in the imagination. It is a symbolic entity.

The eigenform is the imagined boundary in the reciprocal relationship of the object (the "It") and the process leading to the object (the process leading to "It"). In the diagram below we have indicated these relationships with respect to the eigenform of nested boxes. Note that the "It" is illustrated as a finite approximation (to the infinite limit) that is sufficient to allow an observer to infer/ perceive the generating process that underlies it.

Just so, an object in the world (cognitive, physical, ideal, etc.) provides a conceptual center for the exploration of a skein of relationships related to its context and to the processes that generate it. An object can have varying degrees of reality, just as an eigenform does. If we take the suggestion to heart that objects are tokens for eigenbehaviors, then an object in itself is an entity, participating in a network of interactions, taking on its apparent solidity and stability from these interactions.

An object is an amphibian between the symbolic and imaginary world of the mind and the complex world of personal experience. The object, when viewed as a process, is a dialogue between these worlds. The object, when seen as a sign for itself, or in and of itself, is imaginary.

Why are objects apparently solid? Of course you cannot walk through a brick wall even if you think about it differently. I do not mean apparent in the sense of thought alone. I mean apparent in the sense of appearance. The wall appears solid to me because of the actions that I can perform. The wall is quite transparent to a neutrino, and will not even be an eigenform for that neutrino.

This example shows quite sharply how the nature of an object is entailed in the properties of its observer.

The eigenform model can be expressed in quite abstract and general terms. Suppose that we are given a recursion (not necessarily numerical) with the equation

\[
X(t + 1) = F(X(t)).
\]

Here \( X(t) \) denotes the condition of observation at time \( t \). \( X(t) \) could be as simple as a set of nested boxes, or as complex as the entire configuration of your body in relation to the known universe at time \( t \). Then \( F(X(t)) \) denotes the result of applying the operations symbolized by \( F \) to the condition at time \( t \). You could, for simplicity, assume that \( F \) is independent of time. Time independence of the recursion \( F \) will give us simple answers and we can later discuss what will happen if the actions depend upon the time. In the time-independent case we can write

\[
J = F(F(\ldots)))
\]

– the infinite concatenation of \( F \) upon itself. Then

\[
F(J) = J
\]

since adding one more \( F \) to the concatenation changes nothing.

Thus \( J \), the infinite concatenation of the operation upon itself leads to a fixed point for \( F \). \( J \) is said to be the eigenform for the recursion \( F \). We see that every recursion has an eigenform. Every recursion has an (imaginary) fixed point.

We end this section with one more example. This is the eigenform of the Koch fractal (Mandelbrot 1982). In this case one can write symbolically the eigenform equation

\[
K = K \{ KK \} K
\]

to indicate that the Koch Fractal reenters its own indicational space four times (that is, it is made up of four copies of itself, each one-third the size of the original. The curly brackets in the center of this equation refer to the fact that the two middle copies within the fractal are inclined with respect to one another and with respect to the two outer copies. In the figure below we show the geometric configuration of the reentry.
In this example we can denote the initial forms by \( C \) (for circle) and \( B \) (for box). We have \( C \to B \) and \( B \to C \). The rule of imaging is (symbolically):

- If \( P \to Q \) then \( P \to QP \).
- If \( P \leftarrow Q \), then \( PQ \leftarrow Q \).

We start with the mutual reference \( C \leftrightarrow B \). This condition of mutual mirroring can be described by two operators \( C \) and \( B \):

\[
C(P) = CP \quad \text{corresponds to} \quad C \to P.
\]

\[
B(Q) = BQ \quad \text{corresponds to} \quad Q \leftarrow B.
\]

Solving the eigenform equations

\[
\begin{align*}
C(Y) &= X, \\
B(X) &= Y,
\end{align*}
\]

we have the mirror-mirror solution

\[
X = BCBCBCBC\ldots, \\
Y = BCBCBCBC\ldots,
\]

just as in the figure.

We are quite familiar with this form of mutual mirroring in the physical realm where one can have two facing mirrors, and in the realm of human relations where the complexity of exchange (mutual mirroring) between two individuals leads to the eigenform of their relationship.

### 5. Boolean self-reference and the work of Vladimir Lefebvre

Vladimir Lefebvre (1982) models ethical situations involving multiple reflections (I think about your thoughts about me, while you think about my thoughts about you.) using Boolean algebra and a graphical formalism.

The crux of this endeavor begins with examining self-referential equations in the Boolean context. In this context we have the arithmetic of 0 and 1 with \( 1 + 1 = 1 \) and otherwise 0 and 1 behaving as in ordinary arithmetic, with complementation \( (a \to a') \) interchanging them: \( 0' = 1 \) and \( 1' = 0 \).

In interpreting Boolean algebra for logic, we take \( a + b \) to mean "\( a \) or \( b \)" and \( ab \) to mean "\( a \) and \( b \)".

Thus we take \( 1 \) as \( T \) (True) and \( 0 \) as \( F \) (False).

Note that \( a > b \) (\( a \) implies \( b \)) is represented by \( a' + b \) in this system.
A very simple form of Boolean self-reference is the equation

\[ x = x. \]

This just says that \( x \) is equal to \( x \). It is like the biblical “I am that I am.”

A diabolical form of Boolean self-reference is the equation

\[ x = x'. \]

This says that \( x \) is equal to not \( x \), and can be interpreted as the statement of the liar who asserts that he is lying, “I am a liar.” I like to think of the solution to this equation as an oscillation between 0 and 1.

After all, if \( x = 0 \), then \( x = x' = 0' = 1 \) and if \( x = 1 \), then \( x = x' = 1' = 0 \).

So \( x \) oscillates just like a buzzer of a doorbell.

The simplest general form of a Boolean self-referential equation is

\[ x = ax + bx'. \]

What are the possibilities?

We can have

1. \( x = x + x' = 1 \) (constantly true),
2. \( x = 1x + 0x' = x \) (self-affirming),
3. \( x = 0x + 1x' = x' \) (self-denying),
4. \( x = 0x + 0x' = 0 \) (just false).

Thus there does not seem to be a lot of structure in this simplest version of self-reference. However, we should think a bit further and realize that \( a \) and \( b \) can be propositions that have relative truth values and we may not need to know the actual truth values of \( a \) and \( b \). Consider the equation

\[ x = (b' + cx) + cx'. \]

If \( x = 0 \), then we have \( 0 = c \).

So we conclude that for \( x = 0 \) to be a solution, this equation reduces to \( x = x \) or \( x = 0 \). However, if \( x = 1 \) is a solution, then we have \( 1 = b' + c \) and the equation will have a solution just so long as \( b \) implies \( c \) is true. In either case the equation has a non-oscillatory solution. This is the form of the self-referential equation at the base of Lefebvre's analysis of ethics and reflectivity.

The next thing to notice is that \( (a > b) > c = (a + b') + c = ab' + c = (b' + c)a + ca' \).

Thus we have

\[ a = (a > b) > c \]

as an allowable self-referential Boolean equation.

Lefebvre interprets the right hand side of this equation as “\( c \) thinks of \( b \) thinking of \( a \).” Thus the self-reference is “\( a \) is thinking of \( c \) thinking of \( b \) thinking of \( a \).”

Lefebvre takes \( ba \) as notation for “\( a \) implies \( b \).” Thus

\[ b^a = b + a' \]

and \( ba \) stands for \( a > b \), which is interpreted as “\( b \) is thinking about \( a \)” or “\( b \) has an internal image of \( a \).”

Thus our self-referential equation becomes \( a = cx \) where \( x = ba \).

Using Laws of Form

Here is a second take on this theme, using Laws of Form (Spencer 1969) bracket notation. In the Laws of Form notation, we take \( a' = (a) \) and \( ab \) stands for \( a + b \) while the conjunction \( ab \) in Boolean algebra becomes \( \langle a\rangle\langle b\rangle \) in accordance with DeMorgan’s Law. We also have 0 as the void state in Laws of Form and 1 = ( ), the marked state, a single crossing from the void. Then the Boolean arithmetic of 0 and 1 corresponds to the Laws of Calling ( ) = ( ) and Crossing ( ) = “void.”

In Laws of Form notation, “\( a \) implies \( b \)” is written as

\[ \langle a \rangle b = b^a. \]

In Lefebvre’s notation this is the same as \( b(a) = b^a \).

Thus Laws of Form is a useful alternate formalism for this theory.

We can interpret \( b(a) \) as “\( b \) thinks of \( a \).” Consider the self-referential equation

\[ a = \langle a\rangle b c. \]

“\( a \) is thinking of \( b \) thinking of \( c \) who thinks of \( a \).”

This is a self-reference that can be made inside two-valued primary arithmetic, since it never oscillates like \( a = \langle a \rangle \). You can think of this fixed point in the form of the infinite reentry:

\[ a = \langle \langle \langle \langle (\ldots (b) \rangle b) \rangle b) \rangle b) \rangle \ldots c. \]

It is amusing to write this in ordinary Boolean form as \( a = c + da \) where \( d = (b) \) and \( xy = \langle xy \rangle \) and \( x + y \) replaces \( xy \) (LOF juxtaposition). Then we get

\[ a = c + da \]

and so on,

\[ a = c + d^2c + d^3c + d^4c + \ldots \]

\[ a = "c / (1 - d)." \]

The infinite reentry expressions in LOF become an infinite power series in Boolean algebra. This brings us closer to classical mathematics and its role in producing imaginary values.

Lefebvre (1982) in his “Algebra of Conscience” models structures such as

\[ a = \langle b \rangle a \]

\[ b = \langle a \rangle b \]

as “a thinking about himself with an image of \( b \)” and “\( b \) thinking about himself with an image of \( a \).”

We can use the LOF notation to represent the self-referential algebra of Lefebvre, and it is useful to do this.

It is important to see how fixed point equations and reflexivity are intertwined in the Boolean structure. One might think that these concepts would not live in the Boolean context, but of course we do manage to discuss them in the Boolean context of our own thought. So Lefebvre’s model is a microcosm of our condition, and of course this is exactly the point!

This section is just a small introduction to Lefebvre’s theory of reflexivity. It is worth pointing out that he uses the Boolean background skillfully when it is required, but uses the symbolism of reflection on the surface in a way that corresponds to nested linguistic statements. For example, \( a(\langle a \rangle) \) represents “a thinking about \( a \), who has a self-image that corresponds to the true (external) \( a \)” When we evaluate this expression we find

\[ a(\langle a \rangle) = a(\langle \rangle) = a. \]

Thus the non-self-doubting \( a \) is simply himself.

On the other hand, \( a(\langle a \rangle) \) represents \( a \) with an image of himself whose image of himself is false \( \langle a \rangle \). Evaluating this expression, we find

\[ a(\langle a \rangle) = a(a a) = a(\langle a \rangle) = \langle \rangle. \]

Thus the individual with a doubting self image receives a marked value for his skepti-
cism. What about an individual who directly doubts himself? Then we have \( a \left( a(a) \right) = a a = a \). He is in the same boat as the individual with a self-image who doubts. From these examples, we see that the Lefebvre system needs to be examined carefully for its internal meanings. This will be the subject of another paper.

6. Reflexive domains and the magma

A reflexive domain \( D \) is an arena where actions and processes that transform the domain can also be seen as the elements that compose the domain. Every element of the domain can be seen as a transformation of the domain to itself.

In actual practice, an element of a domain may be a person or company (collective of persons) or a physical object or mechanism that is seen to be in action. In actual practice we must note that what are regarded as objects or entities depends upon the way in which observers inside or outside the domain divide their worlds.

It is very difficult to make a detailed mathematical model of such situations. Each actor is an actor in more than one play. His actions undergo separate but related interpretations, depending upon the others with whom he interacts. Mutual feedback of a multiplicity of ongoing processes is not easily described in the Platonic terms of pure mathematics.

Nevertheless, we take as a general principle for a mathematical model that \( D \) is a certain set (possibly evolving in time), and we let \([D, D]\) denote a selected collection of mappings from \( D \) to \( D \). An element \( F \) of \([D, D]\) is a mapping \( F : D \rightarrow D \).

We shall assume that there is a one-to-one correspondence

\[
I: D \rightarrow [D, D].
\]

This is the assumption of reflexivity. Every element of the reflexive domain is a transformation of that domain. Each denizen of the reflexive domain has a dual role of actor and actant.

Given an element \( g \) in \( D \), \( I(g); D \rightarrow D \) is a mapping from \( D \) to \( D \), and for every mapping \( F: D \rightarrow D \), there is an element \( g \) in \( D \) such that \( I(g) = F \). The reflexive domain embodies a perfect correspondence between actions and entities that are the recipients of these actions. See Dana Scott (1980) for a specific construction of reflexive domains relevant to computer science and logic.

An important precursor to this notion of reflexive domain in mathematics is the notion of Gödel numbering of texts. One chooses a method to encode a text as a specific natural number (a certain product of prime powers). Then texts that speak about numbers can, in principle, speak about other texts and even about themselves. If a text is seen as a transformation on the field of numbers, then that text is itself a number (its Gödelian code) and so can be transforming itself. The precision of this idea enabled Gödel to construct mathematical systems that could talk about their own properties without contradiction and he showed that all sufficiently rich mathematical systems have this property. In this way, these systems become self-limiting due to the possibility of statements whose coded meaning becomes “This statement has no proof in the system of mathematics in which it is written,” while the surface meaning of the same statement is a discussion of the properties of certain numerical relations. The domain of numerical relations appears innocuous, and yet it sows the seeds of its own limitations through this ability to reflect itself through the mirror of the Gödelian coding.

The Gödelian example is not just a piece of mathematics. It is a reflection with mathematical precision of the condition of our language, thought and action. We are always equipped to comment on our own doings and in so doing to create new language about our old language and new language about our worlds. All our apparent well-thought-out and directed actions in worlds that seem to extend outward from us in an objective way are fraught with the circularity not just of our meta-comments, but also with the circular return of the consequences of those actions and the influence of our very theories of the world on the properties of that world itself.

We now prove a fundamental theorem about reflexive domains.

We show that every mapping \( F : D \rightarrow D \) has a fixed point \( p \), and an element \( p \) in \( D \) such that \( F(p) = p \). What does this mean? It means that there is another way, in a reflexive domain, to associate a point to a transformation. The point can be seen as the fixed point of a transformation and in that way, the points of the domain disappear into the self-referential nature of the transformations.

Let me tender persuasions. Suppose that \( p = F(p) \). Then we can regard this equation as an expression of \( p \) in terms of \( F \) and itself and write

\[
p = F(p)
\]

and continue in this fashion until the appearance of \( p \) on the right hand side is lost in the depths of the composition of \( F \) upon itself.

The infinite composition of \( F \) upon itself is invariant under one more composition with \( F \) and so \( F(p) = p \) is consistent with this process.

To show that an entity \( p \) is a fixed point for a process \( F \) is to show that \( p \) can be formally identified with the infinite concatenation of \( F \) upon itself. This is an image of the way objects become tokens for eigenbehaviors, in the language of Heinz von Foerster.

Here we show that eigenforms exist in reflexive domains without an infinite limit. The interested reader should compare this argument with the work of William Lawvere (1972). Lawvere proves a more general result in the context of Cartesian closed categories. We have taken his argument and shaped it particularly for this discussion of reflexivity.

**Fixed Point Theorem.** Let \( D \) be a reflexive domain with 1–1 correspondence \( F: D \rightarrow [D, D] \). Then every \( F \) in \([D, D]\) has a fixed point. That is, there exists a \( p \) in \( D \) such that \( F(p) = p \).

**Proof.** Define \( G: D \rightarrow [D, D] \) by the equation \( Gx = F(I(x)x) \) for each \( x \) in \( D \).

Since \( I: D \rightarrow [D, D] \) is a 1–1 correspondence, we know that \( G = I(g) \) for some \( g \) in \( D \).

Hence \( Gx = I(g)x = F(I(x)x) \) for all \( x \) in \( D \).

Therefore, letting \( x = g \), \( I(g)g = F(I(g)g) \) and so \( p = I(g)g \) is a fixed point for \( F \).

Q.E.D.

We shall discuss this proof and its meaning right now in a series of remarks, and later in the paper in regard to examples that will be constructed.
When we define a new element $g$ of $D$ via $gx = F(x)$ for any mapping $F : D \rightarrow D$, and we have a notion of the combination of elements of $D$, $a, b \rightarrow ab$, then we can define $g = g(\cdot x)$ and so get $gg = F(gg)$. Here we have not made a big separation between the elements of $D$ and the mappings, since each element $g$ of $D$ gives the mapping $I(g)x = gx$. But in fact, we could define $ab = I(a)I(b)$ in a reflexive domain.

Whenever anyone comes up with a transformation, we make that transformation into an element of the domain by the definition $gx = F(x)$. We transmute verbs to nouns. The reflexive domain evolves.

The space is not given a priori. The space evolves in relation to actions and definitions. The road unfolds before us as we travel.

Remark 3

We create languages for evolving concepts. The outer reaches of set theory (and category theory) lead to clear concepts, but these concepts are not themselves sets or categories. A good example is the famous Russellian concept of sets that are not members of themselves. Russell’s concept is not a set. Another example is the concept of set itself. There is no set that is the set of all sets.

This very limitation on the notion of a set is its opening. It shows us that set theory can be an evolving language. Language and concepts expand in time.

Here is a transformation on sets: $F(X) = \{x \in X \mid x \neq F(x)\}$. The transform of a set $X$ is the singleton set whose member is $X$. If $X$ is not a member of itself, then $F(X)$ is also not a member of itself. But a fixed point of the transformation $F$ is an entity $U$ such that $U = \{x \in X \mid x = F(x)\}$. We have shown that within the domain of sets that are not members of themselves, there is no fixed point for the transformation $X \rightarrow \{X\}$. This fragment of set theory (sets that are not members of themselves) is not yet a reflexive domain. We shall allow sets that are members of themselves if we wish to have a set theory with reflexivity.

Remark 4: Transcendence

The leap to infinity via self-reference, the production of the finite base of a new level of infinity, the completion of an incompleteness, the emergence of eternity from the world of time – all these metaphors are intimately related to the going back and forth between a process and its eigenform.
Define \( a \ast a = a, \) \( b \ast b = b \) and \( c \ast c = c \). And define \( a \ast b = c = b \ast a, a \ast c = b = \ast a \) and \( b \ast c = a = \ast b \).

In other words, each element combines with itself to produce itself, and any pair of distinct elements combine to produce the remaining element that is different from either of them. The reader can verify that \( \text{TRI} \) is indeed a magma. For example,

\[
\begin{align*}
(a \ast (b \ast c)) &= a \ast (a) = a \\
((a \ast b) \ast (a \ast c)) &= (c) \ast (b) = a.
\end{align*}
\]

Note also that the multiplication in this magma is not associative:

\[
\begin{align*}
(a \ast (a \ast b)) &= a \ast (a) = b \\
((a \ast a) \ast b) &= a \ast (b) = c.
\end{align*}
\]

We will return to this magma in the next section and see that \( \text{TRI} \) is intimately related to the simplest knot, the trefoil knot.

Another example to think about is \( \text{TRI} \), the free magma generated from one element and all its powers; this will place us in the position of creating a rich domain. The examples of magmas related to knot theory, given in the previous section, show that magmas are not just abstract structures, but are related directly to the properties of space and topology in the worlds of communication and perception in which we live.

### 7. Knot sets, topological eigenforms and the left-distributive magma

We shall use knot and link diagrams to represent sets. More about this point of view can be found in the author’s paper “Knot Logic” (Kauffman 1995). In this notation the eigenset \( \Omega \) satisfying the equation

\[
\Omega = \{ \Omega \}
\]

is a topological curl. If you travel along the curl you can start as a member and find that after a while you have become the container.

Further travel takes you back to being a member in an infinite round. In the topological realm, \( \Omega \) does not have any associated paradox. This section is intended as an introduction to the idea of topological eigenforms, a subject that we shall develop more fully elsewhere.

Set theory is about an asymmetric relation called membership.

We write \( a \in S \) to say that \( a \) is a member of the set \( S \). In this section we shall diagram the membership relation as follows:

\[
\begin{array}{c|c|c}
| & a & b \\
\hline
b & a & \{a \} \\
\hline
\end{array}
\]

This is knot-set notation.

In this notation, if \( b \) goes once under \( a \), we write \( a = \{ b \} \). If \( b \) goes twice under \( a \), we write \( a = \{ b, b \} \). This means that the “sets” are multi-sets, allowing more than one appearance of a member. For a deeper analysis of the knot-set structure see (Kauffman 1995). This knot-set notation allows us to have sets that are members of themselves,

\[
\Omega = \{ \Omega \}
\]

and sets can be members of each other.

Here a mutual relationship of \( a \) and \( b \) is diagrammed as a topological linking.

\[
\begin{array}{c}
a = \{ b \} \\
b = \{ c \} \\
c = \{ a, a \}
\end{array}
\]
Here are the Borromean Rings. The Rings have the property that if you remove any one of them, then the other two are topologically unlinked. They form a topological tripartite relation. Their knot-set is described by the three equations in the diagram.

Thus we see that this representative knot-set is a “scissors-paper-stone” pattern. Each component of the Rings lies over one other component, in a cyclic pattern.

Remark. The connection between this formalism and epistemic logic (Hintikka 1962) should be further explored. In epistemic logic the basic expressions are of the form KaKbp (“a knows that b knows that p”).

One specific thing to explore is the problem of common knowledge, which can only be reduced to an infinite number of K’s as in

\[ E = KaKbKaKbKaKb\ldots \]
denoting that “a knows that b knows that a knows that…”

We can write this as

\[ E = KaF \]
\[ F = KbE \]

indicating that

\[ E = "a\ knows\ F" \]
\[ F = "b\ knows\ E" \]

Together these statements indicate common knowledge or mutuality for a and b. Conversely, we can take the linked sets A = [B] and B = [A] as a statement of common knowledge.

Another avenue that should be explored is the relationship between knot set theory and Aczel’s theory of self-referential and non-wellfounded sets (Aczel 1988) and the related treatment by Barwise and Moss (1996).

**Quandles and colorings of knot diagrams**

There is an approach to studying knots and links that is very close to our knot sets, but starts from a rather different premise.

In this approach each arc of the diagram receives a label or “color.” An arc of the diagram is a continuous curve in the diagram that starts at one undercrossing and ends at another undercrossing. For example, the trefoil diagram in the following illustration has three arcs.

Each arc corresponds to an element of a “Trefoil Color Algebra” IQ(T), where T denotes the trefoil knot. The algebra is generated by colors a, b and c with the relations

\[ a * a = a, \]
\[ b * b = b, \]
\[ c * c = c, \]
\[ a * b = b * a = c, \]
\[ b * c = c * b = a, \]
\[ a * c = c * a = b. \]

Each of these relations in the diagram above is a description of one of the crossings in T. The full set of relations describes the coloring rules for an algebra that contains these relations and allows any two elements to be combined to a third element. This three-element algebra is particularly simple. If two colors are different, they combine to form the remaining third color. If two colors are the same, they combine to form the same color.

When we take an algebra of this sort, we want its coloring structure to be invariant under the Reidemeister moves (illustrated below).

This means that when you make a new diagram from the old diagram by a topological move, the resulting new diagram inherits a unique coloring from the old diagram. Then one can see from this that the trefoil must be knotted since all diagrams topologically equivalent to it carry three colors, while an unknotted diagram can carry only one color.

As the next diagram shows, invariance of the coloring rules under the Reidemeister moves implies the following global relations on the algebra:

\[ x * x = x \]
\[ (x * y) * y = x \]
\[ (x * y) * z = (x * z) * (y * z) \]

for any x, y and z in the algebra (set of colors) IQ(T).

An algebra that satisfies these rules is called an Involutory Quandle (Kauffman 1995), hence the initials IQ. Perhaps the most remarkable property of the quandle is its right-distributive law corresponding to the third Reidemeister move, as illustrated below. The reader will be interested to observe that in a multiplicative law QQ, the following operation satisfies all the axioms for the quandle:

\[ g * h = hg^{-1}h. \]

In an additive and commutative version of this axiom we can write \[ a * b = 2b - a. \] Here the models that are most useful to the knot theorist are to take a and b to be elements of the integers \( Z \) or elements of the modular number system \( Z / dZ = Zd \) for some appropriate modulus \( d \). The knot being analyzed restricts the modular possibilities. In the case of the trefoil knot the only possibility is \( d = 3 \), and in the case of the Figure Eight knot (shown after the Reidemeister moves below) the only possibility is \( d = 5 \).

This analysis then shows that there cannot be any sequence of Reidemeister moves connecting the Trefoil and the Figure Eight. They are distinct knot types.
Here is the example for the Figure Eight Knot.

![Figure Eight Knot Diagram]

\[ Z / SZ = \{0, 1, 2, 3, 4\} \] with \( 0 = 5 \)

We have shown how to label the arcs of the knot according to the quandle rule

\[ c = 2b - a = a * b \]

\[ a * b = 2b - a \]

A simple fixed point can be calculated for a given element \( F \) of the quandle.

\[ G(x) = (x * x) * F \]

for a given element \( F \) of the quandle.

Is it then the case that \((x * x) * F = x * g\) for some \( g \) in the quandle?

The answer is, yes, but for a very simple reason:

We have \((x * x) = x\) so that \((x * x) * F = x * F\) and consequently \((F * F) * F = F * F\). In fact, \(F * F = F\), so \(F\) is already its own fixed point. We therefore see that in a quandle the fixed point theorem is satisfied automatically due to the axiom \(x * x = x\) for all \( x \).

On the other hand, if \( F : Q \rightarrow Q \) is an arbitrary mapping from \( Q \) to \( Q \), then we cannot expect that \( F\) will have a fixed point. For example, in the trefoil quandle \( TRI \), suppose we define \( F(a) = b, F(b) = c \) and \( F(c) = a \). Then \( F\) has no fixed point. Note that \( F\) is a structure-preserving mapping. (In this case the composition of \( F\) with itself three times fixes everything. If we make transformations that are permutations of finite sets, then they may be fixed-point free, but some powers of them will certainly have fixed points.)

We have \( F(x * y) = F(x) * F(y) \) for all \( x \) and \( y \) in \( TRI \). For example, \( F(a * b) = F(c) = a = b * c = F(a) * F(b) \).

In order to extend \( TRI \) to a reflexive (right-distributive) magma we would have to add an element \( f \) to the algebra such that \( x * f = F(x) \) for each \( x \) in \( TRI \), take the consequences of that and continue. We leave the exploration of this extension to the reader.

### Reflexivity and Eigenform

We have written the quandle as a right-distributive structure with invertible elements. It is mathematically equivalent to use the formalism of a left distributive operation. In left distributive formalism we have \( A * (b * c) = (A * b) * (A * c) \). This corresponds exactly to the interpretation that each element \( A \) in \( Q \) is a mapping of \( Q \) to \( Q \) where the mapping \( A[x] = A * x \) is a structure-preserving mapping from \( Q \) to \( Q \).


We can ask of a domain that every element of the domain itself is a structure-preserving mapping of that domain. This is very similar to the requirement of reflexivity and, as we have seen in the case of quandles, can often be realized for small structures such as the Trefoil quandle.

As said before, a magma is a domain \( M \) with an operation \( * \) that is left distributive. Magmas are more general than the link diagrammatic quandles. We take only the analog of the third Reidemeister move and do not assume any other axioms \( A \) magma with no other relations than left-distributivity is called a free magma.

The search for structure-preserving mappings can occur in rarefied contexts. See, for example, the work of Laver and Dehornoy (2000; Kauffman 1995), who studied mappings of set theory to itself that would preserve all definable structure in the theory. Dehornoy realized that many of the problems he studied in relation to set theory were accessible in more concrete ways via the use of knots and braids. Thus the knots and braids become a language for understanding the formal properties of self-embedded structures.

Structure-preserving mappings of set theory must begin as the identity mapping since the relations of sets are quite rigid at the beginning. (You would not be able to map an empty set to a set that was not empty for example, and so the empty set would have to go to itself.) The existence of non-trivial structure-preserving mappings of set theory questions the boundaries of definability and involves the postulation of sets of very large size. See Pechcincska (2005) for a good exposition of the philosophical issues about such embeddings and for an approach to wholeness in physics that is based on these ideas.
It is worth making a remark here about sets. Consider the collection Aleph of all sets whose members are themselves sets and such that any investigation into membership will just reveal more sets as members. Typical elements of Aleph are the empty set {}, the set whose member is the empty set {{ }}, and of course various curious constructs that have infinitely many members such as { { { { { } } } } }, { { { { { } } } } }, { { { { { } } } } }, { { { { { } } } } }, ... } and we may even consider sets that are members of themselves (eigen-sets!) such as { { { { { } } } } }.

The key thing to understand about Aleph as a class of sets is that any member of Aleph is, by definition, a subset of Aleph. And any subset of Aleph is, by definition, a member of Aleph. This is a beautiful property of the class Aleph, and it is a paradoxical property if we imagine that Aleph is a set! For if Aleph is a set, then we have just shown that Aleph is in 1–1 correspondence with the set of subsets P(Aleph) of Aleph. If X is any set then we denote the set of subsets of X by P(X). Cantor’s Theorem (proved here in Section 8 and related in that section to the fixed point theory of reflexive domains) tells us that for any set X, P(X) is larger than X.

This means that there cannot be a 1–1 correspondence between Aleph and P(Aleph) if Aleph is a set.

We can only conclude that Aleph is not a set. It is a class, to give it a name. It is an unbroken wholeness whose particularities we can always consider, but whose totality will always elude us. The way that the totality of Aleph eludes us is right before our eyes. Any particular element of Aleph is a set and is a collection of sets as well. But we cannot complete Aleph. Any attempt to approximate Aleph as a set will always have some subsets that have not been tallied inside itself and so the set of subsets of the approximation will grow beyond that approximation to a new and larger domain of sets. Philosophically, this observation of the unreachability of Aleph, the set of all sets, as a set itself is very interesting and important. We see here how a perfectly clear mathematical concept may always remain outside the bounds of the formalities to which it refers and yet that concept is indeed composed of these formalities. It is the leading presence of the ultimately huge and unattainable Aleph that leads us to consider exceeding large sets in the pursuit of a flexibility in the self-embeddings of set theory. At the end of Section 8, we take an alternative view of Aleph and consider what would have to change if Aleph were admitted to be a set.

Enough said about the abstract reaches of the magma.

We should not expect that any given structure is a reflexive space. But it is possible to create languages that can expand indefinitely and thus partake of the ideal of reflexivity.

8. Church and Curry

In this section we point out how the notion of a reflexive domain first appeared in the work of Alonzo Church and Haskell Curry (Barendregt 1984) in the 1930s. This method is commonly called the “lambda calculus.”

The key to lambda calculus is the construction of a self-reflexive language, a language that can refer and operate upon itself. In this way eigenforms can be woven into the context of languages that are their own metalanguages, hence into the context of natural language and observing systems.

In the Church-Curry language (the lambda calculus), there are two basic rules:

1. Naming. If you have an expression in the symbols in lambda calculus then there is always a single word in the language that encodes this expression. The application of this word has the same effect as the application of the expression itself.

2. Reflexivity. Given any two words, A and B, in the lambda calculus, there is permission to form their concatenation AB, with the interpretation that A operates upon or qualifies B. In this way, every word in the lambda calculus is both an operator and an operand. The calculus is inherently self-reflexive.

Here is an example. Let GA denote the process that creates two copies of A and puts them in a box.

\[ GA = \text{AA} \]

In lambda calculus we are allowed to apply G to itself. The result is two copies of G next to one another, inside the box.

\[ GG = GG \]

This equation about GG exhibits GG directly as a solution to the eigenform equation

\[ X = X \]

thus producing the eigenform without an infinite limiting process.

More generally, we wish to find the eigenform for a process F. We want to find a J so that F(J) = J. We create an operator G with the property that

\[ GX = F(GX) \]

for any X. When G operates on X, G makes a duplicate of X and allows X to act on its duplicate. Now comes the kicker.

Let G act on itself and look!

\[ GG = F(GG) \]

So GG is a fixed point for F.

We have solved the eigenform problem without the excursion to infinity. If you reflect on this magic trick of Church and Curry you will see that it has come directly from the postulates of Naming and Reflexivity that we have discussed above. These notions, that there should be a name for everything, and that words can be applied to the description and production of other words, allow language to refer to itself and to produce itself from itself. The Church-Curry construction was devised for mathematical logic, but it is fundamental to the logic of logic, the linguistics of linguistics and the cybernetics of cybernetics.

An eigenform must be placed in a context in order for it to have human meaning. The struggle on the mathematical side is to control recursions, bending them to desired ends. The struggle on the human side is to cognize a world sensibly and to communicate well and effectively with others. For each of us, there is a continual manufacture of eigenforms (tokens for eigenbehavior). Such tokens will not pass as the currency of communication unless we achieve mutuality as well. Mutuality itself is a higher eigenform. As with all eigenforms, the abstract version exists. Realization happens over the course of time.
9. Cantor’s diagonal argument and Russell’s paradox

Let \( AB \) mean that \( B \) is a member of \( A \).

**Cantor’s Theorem.** Let \( S \) be any set (\( S \) can be finite or infinite).

Let \( P(S) \) be the set of subsets of \( S \). Then \( P(S) \) is bigger than \( S \) in the sense that for any mapping \( F: S \to P(S) \) there will be subsets \( C \) of \( S \) (hence elements of \( F(S) \)) that are not of the form \( F(a) \) for any \( a \) in \( S \). In short, the power set \( P(S) \) of any set \( S \) is larger than \( S \).

**Proof.** Suppose that you were given a way to associate to each element \( x \) of a set \( S \) a subset \( F(x) \) of \( S \). Then we can ask whether \( x \) is a member of \( F(x) \). Either it is or it is not. So let us form the set of all \( x \) such that \( x \) is not a member of \( F(x) \). Call this new set \( C \). We have the defining equation for \( C \):

\[
C_x = \neg F(x)x.
\]

Is \( C = F(a) \) for some \( a \) in \( S \)?

If \( C = F(a) \) then for all \( x \) we have \( F(a)x = \neg F(x)x \).

Take \( x = a \). Then \( F(a)a = \neg F(a)a \).

This says that \( a \) is a member of \( F(a) \) if and only if \( a \) is not a member of \( F(a) \). This shows that indeed \( C \) cannot be of the form \( F(a) \), and we have proved Cantor’s Theorem that the set of subsets of a set is always larger than the set itself.

Q.E.D.

Note the problem that the assumption that \( C = F(a) \) gave us.

If \( C = F(a) \), then \( F(a)a = \neg F(a)a \). We would have a fixed point for negation. But there is no fixed point for negation in classical logic!

If we had enlarged the truth set to \( \{ T, F, I \} \)

where \( -I = I \) is an eigenform for negation, then \( F(a)a \) would have value \( I \). What does this mean? It means that the index \( a \) of the corresponding set \( F(a) \) would have an oscillating membership value. The element \( a \) would be like Groucho Marx, who declared that he would not join any club that would have him as a member. We would be propelled into sets that vary in time.

Note that our proof of Cantor’s Theorem has exactly the same form as our earlier proof of the existence of fixed points for a reflexive space. The mapping \( FX \to P(X) \) takes the role of the \( 1 \to 1 \) correspondence between \( D \) and \( [D, D] \). The reader will enjoy thinking about this analogy. In the Cantor Theorem we have used the non-existence of a fixed point for negation to deduce a difference between set \( X \) and its power set \( P(X) \). In the study of a reflexive domain we have shown the existence of fixed points, but we have seen that such domains must be open to new elements and new transformations.

Note also how close Cantor’s Theorem is to Russell’s famous paradox.

Russell devised the set \( R \) defined by the equation

\[
R x = \neg xx.
\]

An element \( x \) is a member of the Russell set \( R \) if and only if \( x \) is not a member of itself. To see the contradiction, substitute \( R \) for \( x \) and get

\[
RR = \neg RR.
\]

This appearance of an eigenform for negation tells us that we either must concede temporality to Russell’s construction \( R \), or else banish it from the world of sets.

10. The secret

What is the simplest language that is capable of self-reference?

We are all familiar with the abilities of natural language to refer to itself. Why this very sentence is an example of self-referentiality. The American dollar bill declares, “This bill is legal tender.” The sentence that you are now reading declares that you, the reader, are complicit in its own act of reference. But what is the simplest language that can refer to itself?

The simplest language would have a simple alphabet. Let us say it has only the letter \( R \).

The words in this language will be all strings of \( R \)’s. Call the language \( LS \). The words in \( LS \) are the following:

\[
R, \quad RR, \quad RRR, \quad RRRR, \quad \text{and so on.}
\]

Two words are equal if they have the same number of letter \( R \)’s.

Each word makes a meaningful statement of reference via the rule:

If \( X \) is a word in \( LS \), then \( RX \) refers to \( XX \).

\( RX \) refers to \( XX \), the repetition of \( X \).

Thus \( RRR \) refers to \( RRRR \) (not to itself), and \( R \) refers to the empty word.

There is a word in \( LS \) that refers to itself. Can you find it?

Let us see. \( RX \) refers to \( XX \).

So we need \( XX = RX \) if \( RX \) were to refer to \( RX \).

If \( XX = RX \), then \( X = R \).

So we need \( X = R \).

And \( RR \) refers to itself.

The little language \( LS \) looks like a pedantic triviality, but it is actually at the root of reflexivity, Gödel’s incompleteness Theorem, recursion theory, Russell’s paradox and the notion of self-observing and self-referring systems. It seems paradoxical that what looks like a trick of repeating a symbol can be so important. The trick is more than just a trick.

The Russell paradox (see the previous section) continues to act as a mystery at the center of our attempts to relate syntax and semantics. In that center is a little trick of syntactical repetition.

I would like to think that when we eventually discover the true secret of the universe it will turn out to be this simple.

The snake bites its tail. The Universe is constructed in such a way that it can refer to itself. In so doing, the Universe must divide itself into a part that refers and a part to which it refers, a part that sees and a part that is seen.

Let us say that \( R \) is the part that refers and \( U \) is the referent. The divided universe is \( RX \) and \( RX = U \) and \( RX \) refers to \( U \) (itself). Our solution suggests that the Universe divides itself into two identical parts, each of which refers to the universe as a whole. This is \( RR \).

In other words, the universe can pretend that it is two and then let itself refer to the two, and find that it has in the process referred only to the one, that is, itself.

The Universe plays hide and seek with herself, pretending to divide itself into two when it is really only one. And that is the secret of the Universe and that is the universal source of our trick of self-reference.
The world of recursive emergence and creativity

We have repeatedly insisted that a formal fixed point or eigenform is associated with any transformation $T$ in any domain where infinite composition of transformations is possible. Thus we make $E = T(T(T(\ldots))))$ and find that $E = T(E)$. This is the symbolic fixed point that sometimes corresponds to a stability in the original domain of the recursion. We have also seen that one can take a seed $z$ for the recursion and repeatedly form

$$z, T(z), T(T(z)), T(T(T(z))), \ldots$$

in a temporal sequence or recursive process. Then the finite products of this process can exhibit similarity to the infinite eigenform, and they can also exhibit novelty and emergence structure in ways that are most surprising. It is this appearance of creativity and novelty in recursive process that makes reflectivity more than abstract mathematics and more than a philosophical idea.

The purpose of this last section is to exhibit an example involving cellular automata that illustrates these ideas and gives us a platform for thought. In this example, we are using an algorithm that I call “7-Life.” It is a variant of the Life automaton of John H. Conway (Gardner 1970).

Conway’s automaton is governed by the rule B3/S23 which means that a white square in the grid is born ($B$) when it has 3 neighbors and it survives ($S$) when it has exactly 2 or 3 neighbors. Life has the property that there are many intriguing formations and processes, but statistically most configurations die out to a collection of isolated static patterns (still lifes) and oscillating patterns that do not grow and do not interact outside themselves.

We should mention that there are a vast number of different cellular automata. A good start in learning about these structures is the book by Stephen Wolfram (2002). Wolfram’s book concentrates almost entirely on one-dimensional cellular automata and achieves a qualitative classification of the behaviors found in a comprehensive class of the simplest types of line automata. Wolfram finds that a number of these simplest automata are computationally universal in the sense that they can simulate a universal Turing machine on the one dimensional lattice of the automaton. The automata that we are looking at in this section, and 7-Life in particular, are two dimensional and hardly considered by Wolfram in his treatise. Conway’s Life is also not analyzed in Wolfram. In fact, it was shown that Conway Life is Turing universal by Conway and his collaborators prior to the onset of Wolfram’s work in the 1980s. We mention this background and the difference in dimensionalities to give the reader some perspective so that he will not be surprised and wonder “Why is this phenomenon not discussed in Wolfram?” Indeed the phenomenon of the remarkable emergence of complexity from simple algorithms is the theme of Wolfram’s work. This theme plays significantly in all algorithmic mathematics and in all significant studies of cellular automata.

6-Life, defined by the rule B36/S23 is well-known, as a search on the Internet for Life automata will reveal. 6-Life does not have the qualitative self-sustaining properties that are evident in 7-Life, but there are emergent structures there as well. We concentrate here on 7-Life because the long-term self-sustaining interactions of this automaton make it ideal for studies of long term evolution and the emergence of forms.

7-Life has the rule B37/S23 and has many of the properties of Life, plus the phenomenon that many starting configurations grow, self-interact and produce streams of gliders. The gliders are five-square formations (occurring in Life as well) that occur spontaneously and regenerate themselves, appearing to move along diagonal directions in the process. The most striking property of 7-Life is the long term persistence of such self-interacting configurations, growing slowly in complexity over time.

In Figures 1, 2, and 3 we indicate the result of applying the 7-Life algorithm to a simple and not-quite symmetrical starting configuration, shown in Figure 1. In Figure 2 we see the result of 33911 iterations of the process. We now have a galaxy of complex interactions. The small entities radiating away from the galaxy are gliders, as described above, and if a reader were to watch the process using a computer program, he or she would see a teeming, seemingly random mass of activity. Then in Figure 3 we see that

Figure 1: The starting configuration.  
Figure 2: After 33911 iterations.  
Figure 3: After 49281 iterations.
after 49281 iterations something new has emerged. It seems that a highly patterned dragon is emerging from the chaos of the complex process. The tip of this dragon moves forward relentlessly.

The body of the dragon interacts with the glider radiation and begins to roil in the chaotic process. So far, the growing tip of the dragon has not interacted with any gliders.

Figures 4 gives a close-up of the tip of the dragon and Figure 5 isolates the generator, GG, of the dragon itself. This configuration GG of 16 squares in mirror symmetry, when placed on an otherwise blank lattice, will generate the dragon in the 7-Life algorithm.

What has happened is that this 16-square generator GG has appeared in the course of the complex interactions, and it has had enough room to move forward in its own pattern – forming the dragon behind it and periodically regenerating itself. The generator of the dragon, GG, is not our invention. GG is a natural consequence of the complex process of 7-Life. GG emerges, but with much lower probability than the gliders. The result is an appearance of novelty and creativity in the complex process as it happens over time. We can only speculate what more complex entities would eventually emerge in 7-Life over many more iterations.

In the same way, DNA emerges from the complex process of the world of the earth and sun.

We see from this example that eigenforms that are processes, such as the self-generating GG, can and will emerge of their own accord from complex systems based on recursion. In this sense, such systems begin to generate their own reflexive spaces. The novel and self-reproducing forms that emerge from them can be seen in a similar light.

All these observations are made by an observer. The observer is clever only in the distinctions that he or she makes, and that is enough to found an entire universe.

12. Discussion

In this paper we have covered a number of mathematical structures related to the concept of reflexivity. We have defined the notion of a reflexive domain D as a domain where the elements of that domain and the mappings of the domain to itself are in 1–1 correspondence.

In such a context, every object is inherently a process, and the structure of the domain as a whole comes from the relationships whose exploration constitutes the domain. There is no place to hide in a reflexive domain, no fundamental particle, no irreducible object or building block. Any given entity acquires its properties through its relationships with everything else. The sense of such a domain is not at all like the set theoretic notion of collections or unrelated things, or things related by an identifiable property. It is more like a conversation or an improvisation, held up and moving in its own momentum, creating and lifting sound and meaning in the process of its own exchange. Conversations create spaces and events, and these events create further conversations. The worlds appearing from reflexivity are worlds nevertheless, with those properties of partial longevity, emergence of patterns, and emergence of laws that we have come to associate with seemingly objective reality.

Note

1. According to Webster’s New Collegiate Dictionary (1956) a gremlin is “one of the impish foot-high gnomes whimsically blamed by airmen for interfering with motors, instruments, machine guns, etc.; hence any like disruptive elf.”

THE AUTHOR

Louis H. Kauffman has a B.S. in mathematics from MIT and a Ph.D in mathematics from Princeton University. He is Professor of Mathematics at the University of Illinois at Chicago. Kauffman works on topology, knot theory, quantum topology, form and cybernetics. He is the founding editor and editor-in-chief of the journal Knot Theory and Its Ramifications and the editor of the book series on Knots and Everything. He is the author of the column Virtual Logic for the journal Cybernetics and Human Knowing. Kauffman is the author of numerous books and papers on knot theory and is well-known for his discovery of the bracket state sum model of the Jones polynomial and for his discovery of a generalization of the Jones polynomial called the Kauffman polynomial, for his discovery of virtual knot theory and other topological structures. He is the recipient of the 1993 Warren McCulloch award of the American Society for Cybernetics and the 1996 award of the Alternative Natural Philosophy Association for his work on discrete non-commutative electromagnetism. He was president of the American Society for Cybernetics from 2005 to 2008. Kauffman plays the clarinet in the Chicago-based ChickenFat Klezmer Orchestra.

Figure 4: The growing tip.

Figure 5: The generating tip GG.
References


Received: 15 May 2009
Accepted: 7 July 2009

Volume 4 · Number 3 · July 2009

137
Conversation
Possibilities of its Repair and Descent into Discourse and Computation

Klaus Krippendorff  University of Pennsylvania, USA <kkrippendorff@asc.upenn.edu>

Introduction
In my answer to Ernst von Glasersfeld’s (2008) question, “Who conceives of society?”, I proposed a radically social constructivism (Krippendorff 2008a) that overcomes what I perceive to be an unfortunate cognitivism in von Glasersfeld’s, Heinz von Foerster’s, and Humberto Maturana’s work. Since then, I published two other papers on the subject. One (2008b) moves the notion of human agency into the center of my project, focusing on its role in conceptions of social organizations – a concept less grand than “society”; one (2008c) teases out several reflexive turns that have grown in cybernetics but cannot be subsumed by the epistemology of radical constructivism and second-order cybernetics, which privileges observation and a representational theory of language over participation in conversation and cooperative constructions of reality. In all of these efforts, conversation has become the starting point of my conceptualizations of being human. In this essay, I wish to discuss what conversation entails, how it is maintained, and under which conditions it degenerates into something else.

Since Martin Heidegger, many philosophers have based their work on the contention that humans live in language. I concur with this proposition but must warn that there are several conceptions of language (Volosinov 1986) and it is important to be clear about the specific conception of language when subscribing to such a proposition. Linguistic conceptions of language are largely due to Ferdinand de Saussure’s (1916) unfortunate but consequential distinction between “langue” and “parole.” For him, langue, the French word for language, is the relatively enduring system of rules and conventions common to all of its speakers, and parole, the French word for speaking, is what speakers do with language. The latter is considered full of idiosyncrasies, marred by individual incompetencies, entirely situational, messy, difficult to study, and hence excluded from the object that linguistics constructs and calls language. Also, for Saussure, langue and parole are what individuals speak. That we always speak in the expectation of being understood by those addressed, in social relations with others, not merely expressing our experiences to the world – inter-individual relationships – do not enter traditional linguistic inquiries, socio-linguistics nudge excepted. In my view, linguists study a convenient abstraction from processes of conversations, purporting to be the systematic and conventional structure that governs individual speakers. It construes that abstraction as the government of individual speech.

For me, Humberto Maturana and Francisco Varela’s (1980, 1987) term “languaging,” or “the use of language,” brings the linguist abstraction back to where it is embodied, in real people speaking with each other. Languaging is a process of mutual human engagement. It is not just a biological capability. Languaging has a history: developmentally, in the sense that individual humans learn it from each other; etymologically, in the sense that spoken utterances and written words have lineages that go back generations of uses by largely unrecognized cultural ancestors; and ontogenetically, in the sense that it goes hand in glove with the evolution and use of cultural artifacts. Languaging is a social or inter-personal phenomenon, not a cognitive one.

For Ludwig Wittgenstein (1953; Schulte 1992), language is a game we play with each other, and the meaning of its words is the history of acquiring their use. When we learn a language, we learn to coordinate ourselves with present others. This is quite consistent with Maturana’s (1988) conception of language as the con-sensual coordination of consensual coordinations of actions. The dash
between “con” and “sensual” is mine and is intended to highlight the jointly sensing of (focusing on) something and each other by speakers, and to prevent the common reading of “consensual” as relating to consensus or agreement. Playing soccer, for example, requires much coordination among players relative to a moving ball. But what makes handling that ball a soccer game has much to do with the interpretation of written rules, for example, by referees who must declare something to be a violation or a scored goal, or which team won.

I contend that Wittgenstein’s choice of the game metaphor may not have been an entirely happy one as it suggests language as a means of accomplishing something; a tool, for example, to decide who won the game. Surely this is not what he implied. Rather, his language games do not need to be finite and may well be ongoing, a “way of life” in which people have the courage to change their being with each other. I have similar misgiving with the idea of language as the coordination of coordinations of actions. Language does not control anything. Speakers interact with each other and define themselves interactively, not as individual actors, but as participants, acting jointly (Shottter 1993). Even in a soccer game, not all participants are eager to win the game. Besides the two teams of players, including their coaches, there are referees, audiences, field owners, and their employees, whose diverse realities are necessary but not questioned during a game. As Wittgenstein reminds us, using language does something. In the process of speaking, realities are cooperatively created and maintained in which speakers constitutively participate in relation to each other. Human relations, soccer games, cities, and technologies are interactive accomplishments, cognition always playing only a part in them. What individual soccer players have in mind may well affect the outcome of the game but does not determine its end.

In (2008c), I worked towards the conclusion that cybernetics is an interdisciplinary discourse that brings radically reflexive realities into being, which includes attention to a host of familiar constructions from feedback loops, self-references, recursions, autonomies, to its own constructive use of language. There I suggested that second-order cyberneticians do not go far enough when they merely reflect on their observations, taking responsi-
engagement with each other has a history that precedes our participation in it and this history inevitably resonates in ongoing conversations. Conversation is one explanation that constitutes itself in practicing human togetherness.

The following two sections describe conversation from two contrasting positions. The first applies von Foerster and Maturana's variously articulated conception of a standard scientific observer (here of conversation) whose aim is to be conscious of his or her acts of observing and describing his or her observations/constructions to others. The second takes the position of a participant in conversations whose competencies reside in contributing to what is happening there. The difference between these two positions is not found in the difference between objective and subjective accounts of the same phenomena but between outsider and insider accounts. All accounts occur in conversations and are offered in the first position by one observer (of conversations) to a community of other observers, and in the second position by participants in the very process to be accounted for. I am using the second section not only as a critique of the first, showing the epistemological limitations of celebrating observers and observations, but also as a reference to what happens when conversation degenerates into something else.

**Conversation observed**

Morphologically, “con-” means together, joint, or among, and “-versation” has many roots, from making “verse” out of experiences as poets do, being “consens” in a subject matter, to a “version,” translation or interpretation of something, including of reality. The *Oxford English Dictionary* (1991: 868) traces “Conversation” to the 12th century and gives its earliest meaning as “The action of living or having one's being in a place of or among persons” and “The action of consorting or having dealings with others; living together; commerce, intercourse, society, intimacy.” In the 16th century, conversation became “Interchange of thoughts and words; familiar discourse or talk.” This etymology suggests the meaning of conversation to be remarkably stable. Its overriding use as a *way of being together in talk and interaction* serves me well.

Contrasting dialogue with writing, I suggested: Everything said is said in the expectation of being understood by an addressee. Everything heard as being said is taken as having been said by one person to another. Understanding does not need to be mutual and shared, but needs to be complementary in how it is performed (Krippendorff 2009: 159). Minimally, conversation requires two participants in interlacing expectations. Charles Goodwin (1981: 4), citing Erving Goffman (1976), differentiates three listeners to talk. Those who overhear a conversation without being part of it and without the expectation or ability to respond, those who are part of a conversation and (in case of three or more participants) are addressed by the speaker and expected to respond, or those not addressed and not expected to respond. Goffman and Goodwin thought of overhearers as casual bystanders. I am including as bystanders the observers of conversations – for example, through a one-way mirror – the listeners to wire tapped telephone conversations, the viewers of verbal interactions on a movie screen, and, most important here, the conversation analysts, typically working from transcripts of naturally occurring talk. The latter are scientific observers of conversation and I maintain their view is necessarily unlike the view of involved participants.

As a scientific observer, overhearing and recording conversations from their outside, Robert Nofsinger (1991) considers conversations as:

- **Mundane activities** among those observed together. Everyone is able to engage in conversation with others without specialized knowledge, preparation or equipment. This observation may need to be qualified by noting that conversation is learned. Mothers incessantly talk to their babies, initially pleasing to get at least a smile in response. It is not clear how babies or children listen, but in time, their participation becomes richer and entirely natural or mundane. Then Nofsinger’s observation applies.

- **Common occurrences.** Conversations are observed everywhere, at home, at work, while shopping, in public places, on the telephone, and between waking up in the morning and exchanging intimacies with a partner at night. While mostly taking place among acquaintances, conversations also occur among strangers such as when waiting in line for a cashier or in a doctor’s office.

- **Interactively unfolding in time.** Participants take turns and respond to each other’s utterances. A conversation essentially is a sequential activity. It creates its own history. This history can be recorded, videotaped, transcribed, and examined in detail, providing analyzable data.

- **Locally managed.** During the course of a conversation, participants themselves determine who speaks, for how long, and in which order. Responsibility for maintaining a conversation is distributed among those present.

- **Accompanied by other activities.** Participants do not merely say something to each other when they talk. They also do something at the same time. Activities may include non-verbal expressions – gestures, eye contact, variations in voice – but they also establish relationships among speakers and coordinate parallel activities. Conversations between the pilot and copilot direct an airplane’s flight; within a team of designers, result in a novel technology; between therapist and a client, produce new realities, ostensibly for the client but in fact for both; among business partners, shape actionable agreements; or among the employees of a social organization, determine what that organization is and how everyone contributes to it. Conversations coordinate the realities of everyday life.

Other scholars consider conversations as:

- **Extendable to mediated activities.** Although speaking a language is acquired in the bodily presence of others in conversation, once learned, conversations can continue through interactive media, between participants out of sight of each other. Exchanging written letters, once the only form of mediated conversation, is being replaced by telephone conversations, online discussions, email, and texting. While all mediated conversations omit some features of face-to-face conversations – sight in telephone conversations, identity in some text-based internet discussions – they always extend desirable dimensions – distance, for example. Yet, even in mediated conversations, participants are aware of each other.

**Academic interests in conversations** assume conversations to be

- **Analyzable and theorizable,** usually from recordings and transcripts that allow the conversation analyst to examine and
reexamine the data for patterns that may otherwise escape even the most attentive listening, or in the case of mediated conversations, casual reading.

Theories based on such data always are and cannot be anything other than the theories of observers, not of the observed participants – unless the latter articulate their theory in use, which is rare. However, the position of observers and participants should not be confused on epistemological grounds. Also, theories always reflect the disciplinary interests of theorists in a limited aspect of the available data. For example, therapists typically look for clues to a diagnosis of their clients’ mental problems, ignoring everything else, including their own creative contributions to this end. Employers may examine interview data to predict whether an interviewee will fit their job description; cognitive scientists select from verbal interactions that which allows them to infer what is going on in participants’ minds. Conversation analysts are not immune to such limitations either when seeking to invent rules that could explain the organization of talk and exchange of written messages, except that their theories tend not to aim at generalizations but are satisfied with moment-to-moment explanations.

It is often taken for granted that conversation analysts can hardly proceed without speaking the language of the participants in observed conversations, nor can they succeed without conversational experiences on their own. Even the transcripts they prepare are cultural artifacts that speak of the analysts’ competencies to engage in and write down what they observe. Reliance on such data questions the detachment that conversation analysts seek to project in their analyses and explanations.

Insightful analysts may well have been part of the very conversations they subsequently analyze. Goodwin (1981), for example, taped many birthday parties and gatherings among friends, bringing insider experiences into his analysis. But being forced to demonstrate the validity of a conversation analysis in terms of quotes from transcripts or clips from video recordings encourages explanations of sequential interactions, turn taking, and how categories of utterances follow each other. Such sequential actions, turn taking, and how categories of language (Volosinov 1986) that linguists pursue as addressable or the source of their voices are considered unreliable and not born out by observational facts. Preservation of objectivity was one reason for linguistics to exclude parole and conversations from their object of study. Conversation analysts are not committed to the abstract-objectivist notions of language (Volosinov 1986) that linguists pursue but also shy away from becoming conversationally involved, leaving their preferred observer role. In the social sciences, participant accounts largely are considered unreliable and not born out by observational facts. Preservation of objectivity was one reason for linguistics to exclude parole and conversations from their object of study. Conversation analysts are not committed to the abstract-objectivist notions of language (Volosinov 1986) that linguists pursue but also shy away from becoming conversationally involved in their object of analysis.

To appreciate the severe limitations of understanding conversations by overhearing or observing conversations from the outside, let me now describe, as far as possible, conversation from within the process, as a participant.

**Authentic conversation**

In existential philosophy, authenticity has to do with being true to one’s self despite pressures from society to be otherwise. There, authenticity is celebrated as an individualist ideal that denies the conversational reality of being human. I am using authenticity here to refer to the pleasure of participating in togetherness in which one is free to speak for oneself, not in the name of absent others, not under pressure to say things one does not believe in, and not having to hide something for fear of being reprimanded or excluded from further conversations. But I will be more specific than that.

Authentic conversation is not easily, if at all, identifiable from the outside. How would an observer access someone’s construction in progress, why something is said, and what is not being said? Questions of this kind should not be dismissed as being subjective. Inasmuch as participants in conversations can be asked and may be willing to account for their feelings, the act of making them public, where they can be dealt with in the very conversations that elicit them, renders feelings – supposedly subjective – inter-subjectively acceptable. One is reminded of Wittgenstein’s argument against private language. Participant accounts are not only richer in meaning and closer to what is happening inside conversations. Self-organizing systems, by definition, develop their own identities, their own realities, and their own meanings for what occurs within their boundaries. For outsiders, it is extraordinarily difficult, perhaps impossible, to explain why participants say what they say and how a conversation develops the way it does, except for the above-mentioned possibility of asking questions of the participants, in effect intervening in the conversation of interest, thus bringing their own conversational experiences into the very conversation to be analyzed.

By analyzing the transcripts of conversations, conversation analysts notice patterns that may mean nothing to participants inside conversations. To claim that participants in conversation are unaware of the patterns that conversation analysts are “discovering,” or more correctly, “constructing,” is epistemologically untenable – unless analysts step out of their observer role, explore their hypotheses with the participants in a conversation, and thus become conversationally involved, leaving their preferred observer role. In the social sciences, participant accounts largely are considered unreliable and not born out by observational facts. Preservation of objectivity was one reason for linguistics to exclude parole and conversations from their object of study. Conversation analysts are not committed to the abstract-objectivist notions of language (Volosinov 1986) that linguists pursue but also shy away from becoming conversationally involved in their object of analysis.

To appreciate the severe limitations of understanding conversations by overhearing or observing conversations from the outside, let me now describe, as far as possible, conversation from within the process, as a participant.

**Occurring in the presence of addressable and responsive individuals.** In authentic conversations, participants distinguish themselves and each other by the contributions they make to them. The act of distinguishing oneself is public. It does not impose identities on others, which is what observers are destined to do. When participants cannot be seen as addressable or the source of their voices cannot be distinguished – for example, when in a large and anonymous crowd – conversation is no longer authentic.

**Maintaining mutual understanding.** In conversations, mutuality, agreement, and
coordination of understanding and acting are of central concern for all participants. However, since cognition cannot be observed and nobody can compare their own understanding with that of others, in conversations, understanding or the lack of it, is performative and evident in certain speech acts, such as “I understand,” “I agree,” or “tell me more.” Here, acknowledging understanding does not mean similarity or sharing of conceptions, its affirmation constitutes an invitation to go on, including to other subjects.

Observers, by contrast, are effectively excluded from the possibilities of checking their understanding of what they overhear against the performative understanding among participants in conversation. In this respect, analysts of transcripts of conversations or written exchanges are literally “out of the loop,” isolated, and responsive, at most, to their scientific community of equally detached observers.

- Self-organizing and constituted in the contributions their participants make to each other. Conversations are communicational closed. They are not abstracted from anything. They are embodied in real participants’ talking and listening to each other, responding to what they hear, and acting accordingly. The identity of a conversation – dinner conversation, political deliberation therapeutic session, focus group discussion, business meeting, or design project – emerges from talk and text generated within that conversation. With the emergence of conversational identities comes the feeling of being part of it, referring to its participants by the inclusive “we.” How the responsibility to maintain the flow of conversational moves is distributed among participants and the direction in which a conversation is going is always uncertain – save for one’s own contribution. Among participants, this uncertainty is not a deficiency, however. Participants trust each other to make sense of what is said.

Observers who seek to understand a conversation from a recording of what happened, looking at it from a God’s eye view, cannot possibly appreciate the feeling of being part of it, the feeling of being able to shape an always evolving conversation, and the feeling of trusting each other to maintain the flow. As Michael Billig (2006) noted, we have a rich vocabulary of inner processes – feelings, thoughts, attitudes, experiences, memories and reasons – in terms of which psychologists construct the cognitive processes of their interest without being observable. However, it is because the conversational use of these words is public and coordinated with other speakers of a language that they become meaningful in conversations, not as description of individual states but as performing certain speech acts.

- Intuitive, not rule governed. Authentic conversations are embodied practices. Turn taking, topic switching, coordination of reality constructions is natural, requiring no reflection, no preparation, no special training – as Nofsinger said, notwithstanding the fact that children, born into a community, need to learn joining its conversations. Children do not learn rules, however, and then apply them. They learn to interact with others by speaking much like how they see and hear others interacting with them. Authentic conversations do not follow rules; they give birth to further conversations. Only after sufficient conversational competencies are acquired is it possible to talk of improper practices – “do not interrupt,” “don’t be rude,” or “listen!” from which conversational conventions may emerge. But authentic conversations may go on without them.

Conversation theorists may well draw useful distinctions from the transcripts of conversations, for example, by analyzing conversational triples and adjacency pairs, formulating and testing hypotheses about how natural conversations are organized (Goodwin, 1981), postulating conversational maxims (Grice, 1975, 1978), or theorizing a universal pragmatics for ideal speech situations (Habermas, 1970, 2001). But all of these grand theoretical precepts are constructions by and for outsiders to conversations.

Conversation analysts have the tendency to claim that participants implicitly follow the rules they have invented. This claim is epistemologically preposterous, however. Drawing on Sigmund Freud, Billig (2006) makes a useful distinction between the unconscious and the preconscious. The former is an observer’s construction of cognition that is inaccessible to an observed individual (and often related to oppression). The latter is an observer’s construction of what that individual does not attend to at the moment, takes for granted while conversing with others. But from the perspective of social construction, there is the possibility that conversation analytic vocabularies enter a conversation and start coordinating participants’ talk whether of cognitive conditions or conversational rules. In other words, while the results of conversation analysis may not have anything to do with how conversation is practiced, teaching conversation theoretical explanations diverts practitioners’ attention from what they had been doing naturally.

- Dialogically equal. By dialogical equality I mean that every participant in a conversation has the possibility of contributing to it. Nobody feels excluded. Every contribution, even silence, is respected and appropriately responded to.

Indeed, participation is rarely observed equal. Some participants inevitably speak more than others do, leading to claims of observed power inequalities within conversations. Moreover, participants usually have unequal resources (experiences) to contribute. Turn taking is inherently asymmetrical. However, such interpretations of observed differences in frequencies as indicators of inequalities may not matter to insiders to whom unequal experiences may not be detrimental to authentic conversations, more likely, they are what keeps a conversation alive. Even without making an observable contribution, the perception of being able to contribute when the opportunity arises and be accepted for what one says is all that matters. Needless to say, dialogical equality is not observable from outside a conversation. Participants may not notice it either but might articulate its lack.

- Creating possibilities of participation. Conversations may well take place while doing a job. But besides correlations with a purpose, conversations are inherently creative, offering participants possibilities to contribute and realize themselves in the contributions they and others make to the process. One obvious example of opening possibilities of participation is to raise questions not previously answered, inviting addressees to construct answers. Conversational possibilities expand when participants assure each other that their contributions are understood, important, and appreciated (Brown & Levinson 1987), and that their creativity is appreciated. Creating and maintaining possibilities for others relates to von Foerster’s (1981: 308) ethical imperative: “Act always so as to increase the number of choices.” Here, I am embedding his imperative in the context of social interactions. Socially relevant
choices, not their numbers, are the gifts that partners in communication can offer each other (Krippendorff 2009: 34).

Obviously, possibilities can be created, pondered, exhausted, and constrained, but not observed. It should also be noted that not all questions may invite participation, as I shall discuss below.

Irreversible, progressive, and unique. For participants, conversations never repeat themselves. Each turn is experienced as unique; each utterance reveals its speakers’ shifting perspectives. As Heraclitus suggested, “you cannot step in the same river twice.” Participants have numerous conversational moves available to alert each other to redundant threads: “here we go again!” “didn’t you already tell that story,” “old news,” etc. Indeed, it makes no sense to repeat stories unless they have been forgotten or decisions unless they have not been followed up or been previously undone.

For conversation analysts, each transcript may well be unique as well. However, scientific analysis calls for the identification of recurrent patterns and generalizations at the cost of excluding the very uniqueness to which the participants in conversations respond. Observers tend to be blind to the unique contributions made in conversations. Participants tend to be blind to the repetitions they take for granted. Evidently, observers and participants construct realities that are orthogonal to each other but not necessarily incompatible.

Coordinating constructions of reality. Conversations always leave artifacts behind, minimally the memories of their own history. Other artifacts include the always evolving relationships among participants. But most important are the changes that participants introduce into the world while being in and after participating in conversations: decisions with practical consequences, institutionalizations of procedures, projects, designs or texts, and realizations of diverse technologies. Rarely do these artifacts correspond to any one individual’s cognition. Participants supplement each other’s contributions (Gergen 1994). Indeed, furniture, cars, computers, the internet and cities are designed in the course of many conversations, having long histories with changing participants but a common thread. Conceptions of these artifacts need not be shared and mostly cannot be articulated in full by any one individual but may complement each other in the interactions that set these artifacts in motion.

Conversation theorists cannot achieve such coordinations for their theories – unless they join the conversation they are theorizing and become active participants, no longer observers. Similarly, theorists of technology are comfortable in describing the histories of technological developments, but rarely appreciate the multiple conversational grounds of such developments, much less dare to forecast technological developments. The belief in technological determinism is an extreme case of denying the role of language and social interaction that drives such developments.

Continuable in principle. From the perspective of external observers, conversations may be short, such as between occupants of neighboring seats on a city bus, terminating when they no longer sit next to each other, or long, such as between teenage friends who talk for hours on the telephone. For observers, both examples are finite in time. But what they have in common is the possibility of their continuation at a later time, at a different place, and perhaps including new participants, no matter what has happened in between separate encounters. When children move out of their family – for instance, when they go to college – and stay in touch with their family members and friends by telephone, email, or text messaging, they continue to weave the conversational realities they had started long ago, albeit by different means, across geographical distances, and under continuously changing circumstances. Conversations can degenerate into other forms of interactions, incompatible with the above, and, in the extreme, when violence enters, which is a categorically different way of being together.

Evidently, there are vast differences between how participants see themselves in authentic conversations and what outside observers, conversation analysts, can record, analyze, articulate, and theorize. The two positions are con-sensually different, distinguished by unlike epistemologies, unlike relationships to their objects of attention, and unlike experiences with the subject matter of talk or writing. I am not devaluing the position of the conversation analyst, but wish to highlight that their reality constructions necessarily differ from the ones of those who are conversationally involved with each other.

Accountability and possibilities of repairs

The above depicts conversations as self-organizing and unproblematic verbal and non-verbal interactions among participants, including the constructions of reality they produce. Authentic conversation is typical among trusting friends but also among strangers who, having nothing to lose, feel alive in each other’s presence. I do not expect participants able to describe what authentic conversation entails – as I tried above – but to become aware when disruptions of it are experienced.

In everyday life, people do not always respond in perfect alignment to each other. We say things that may not be understood as intended, interrupt someone’s turn, offend someone without wanting to, or talk too much and thereby preempt others from speaking their mind. Besides such unintended disruption of unproblematic interactions, we know of systematic and institutionalized disruptions that we may notice when they occur but fail to address for a variety of reasons. I maintain that conversational competencies include ample possibilities to repair problematic conversational sequences within them. Whether or not we utilize these linguistic resources and how aware we are of these possibilities is a big question that I cannot answer here. Often it is only after encountering the efforts of others to repair problematic conversations that we are made aware of how we deviate from authentic conversation – without implying the ability to articulate just how a conversation got astray. Possibly the most important linguistic resource for repairing disruptions of authentic conversations is accountability.

I contend that everything said is said not only in the expectation of being understood by addressees but also in the expectation of being held accountable for what was said or done. As John Shotter (1984, 1993) suggests, speakers tend to articulate their contributions to a conversation not merely in response to other speakers but also with possible accounts in mind in case their contributions are challenged. The process of holding participants accountable may be initiated by noting an infelicitous, untoward, or problematic conversational move, action, or sequence of exchanges. Expressing dis-ease with some-
one's contribution – sometimes called meta-communication – amounts to a momentary disruption of that flow and implies a request for an account by the presumed source of that dis-ease. Requests for an account may also be made directly: "Why did you say that?" "What do you want to accomplish with that proposal? "Why do you come so late?" The account subsequently given is then evaluated and either accepted or rejected, and in case of the latter, a new account may be requested, until the issue is resolved (Buttny 1993).

The most typical accounts are explanations, justifications, excuses (Mills 1940; Scott & Lyman 1968), and apologies. The interactions they set in motion are part of the conversation. They differ from the unproblematic flow of a conversation by focusing on the interaction in question, not on what they construct.

- **Explanations** are least disruptive of conversations. They respond to assertions such as "I don’t understand," "I am not following you," and questions such as "can you clarify?" or "what do you mean by that?" Explanations, once accepted as making sense, have the effect of coordinating participants' understandings performatively and bringing a conversation back to an unproblematic flow. Good explanations rearticulate or expand what had been said in terms compatible with listeners' background of understanding.

- **Justifications** acknowledge a speaker's agency in an actual or anticipated happening, and respond to expressed doubts of the merit of that happening. Justifications may be defensive when responding to challenges or preparatory when actions are proposed with the intent to seek approval. Often justifications are used to enroll listeners into the speaker's project (Krippendorff 2008b). Once justifications are accepted, conversation can proceed to other topics.

- **Excuses**, by contrast, deny a speaker's or actor's agency, intention, or involvement in what happened and offer grounds for not being responsible for it. Typical excuses are appeals to external causes, lack of knowledge, accidents, being under the influence of drugs, or having acted on the orders of a superior. The latter may shift blame to someone else, which is a common diversion. If accepted, excuses render speakers blameless and enable them to continue their participation in the conversation. Excuses rely on narratives that are intended to be compelling, but not necessarily true by extra-conversational standards. Excuses do not change the condition for which they are offered.

- **Apologies** admit responsibility for an offensive conversational move or action, express regret, and imply the promise not to repeat it in the future. Unlike excuses, apologies admit the actor's agency. Accepting an apology forgives the perpetrators of offensive conversational moves or actions and is a way to continue the conversation in the hope that the offense will not recur.

Shotter's (1984, 1993) observation that all speakers talk in the expectation of being held accountable by listeners for what they say and do applies to the act of giving accounts as well. Accounts too are always articulated in the hope of being accepted and only those that have that chance are offered. Although accounts may well appeal to general conventions – rationality, common benefits, individual values, or established practices – such conventions are effective only in the very conversations in which participants are willing to let them stand. Inasmuch as the mutual acceptance of practices of living together is a matter of ethics by definition, successful accounts provide narratives that participants in conversation consider ethical. Thus, in repairing problematic conversations, conversation-specific ethical narratives are proposed, tested, and accepted, i.e., narratives that participants can live with and find no reason to object to. The ethics that emerges in repaired conversations has two remarkable features. It is rarely generalizable to all conversations – effectively denying their universality; for example, the universal pragmatics of communication proposed by Habermas (1970) – and it cannot be represented by any one observer's or participant's cognitive construction. Conversational ethical realities are performed in conversations or interactively constructed.

Accounts may be personal, "I was angry," informational, "I didn't know that," related to efficiency, "this is all I could afford," ethical, "I didn’t want to hurt her," moral, “everyone does it,” pragmatic, “it worked in the past,” or institutional, “this is the approved procedure.”

Problematic conversations can be considered repaired when they resume their natural flow. However, conversations are not machines that can be fixed by replacing defective parts. Successful repairs have the potential of leaving memorable residues behind, an awareness of what happened and how it was resolved. Such residues may become part of the history of a conversation and direct that conversation's future along paths not taken without prior repairs. Therefore, a history of successful repairs holds the seeds of conventional accounting practices in terms of which future problematic conversational moves may become explained, justified, excused, or apologized for.

Thus, unless the history of repairs is forgotten, repaired conversation may no longer be quite authentic and I would argue this condition to be most common in naturally occurring conversations.

**Degeneration of conversation**

While language always provides ample resources for repairing untoward conversational moves or actions, this is not to say that all disruptions of the flow of interactions are indeed repaired. Not repairing problematic conversations is not limited to children who are in the process of developing accounting competencies. It applies to competent speakers as well. Failing to repair conversations that have turned problematic has two important social consequences. On the one hand, participants who do not hold each other accountable for what they say or do, whether for reasons of expediency or fear of reprisals, grant implicit permission to continue the untoward practice, which can lead to their tacit legitimation. On the other hand, participants who refuse to give adequate accounts when requested of them claim exceptional privileges, in effect, which can lead to the institution of inequalities and violate the dialogical equality that authentic conversation requires.

There may be reasonable and unfortunate conditions for not practicing accountability. Temporarily suspending conversation to get something more important accomplished might be considered reasonable – as long as this suspension is temporary and mutually consented to. Entrapment of one by another – threads of exclusion from a conversation, induction of fear of retribution, and exercising authority – is always unfortunate because acquiescence inevitably creates burdensome
interpersonal relationships that are incompatible with authentic conversation. The unwillingness to repair problematic conversations is the root cause of conversations descending into other forms of interaction, as I shall exemplify below. The results of such degenerations are where conventional sociological abstractions stand – without adequate reflection on their roots in conversations.

There are innumerable many ways a conversation can degenerate into other forms of social interactions. I can offer only a few examples.

**Physical constraints**
The most benign and not entirely social in nature are physical constraints. Conversations become increasingly difficult when noise competes with participants’ ability to listen to each other’s voices, or when the number of participants grows too large for speakers to address individual participants or to distinguish individual voices, for example at mass rallies, political demonstrations, or public performances. In such situations, participants acquire collective identities that divide participants into, say, performers and audiences or demonstrators and police.

**Dialogical inequalities**
Most obviously, authentic conversation degenerates by tolerating dialogical inequalities. Interruptions of a speaker’s turn can happen carelessly, but they also may be part of accepted discourse practices. For example, it is well known that men interrupt women more often than the reverse. Numerous explanations have been suggested, including a prevailing acceptance of patriarchy. More clearly explainable dialogical inequalities occur at board meetings. Authentic conversation among equals disappears as soon as the CEO or a person in charge of the meeting enters. Such situations are often explained in terms of unequal distribution of power. Power, however, is not what superiors have and subordinates lack. It is not measurable by unequal access to material resources but manifests itself in the unwillingness to hold authorities accountable for what they say or do, and, its complement, in the refusal to provide accounts when requested (Krippendorff 2009: 131–155). Power arises when accountability is not exercised and subsequent interactions are tolerated.

Therapists have sometimes been characterized as conversation managers, which makes therapy different from conversation. Managing focus groups or group discussions, for example, by instructing participants to list their ideas on a predefined issue, putting them on public display, and then proceeding to group them gives the impression of dialectic equality by granting every participant a voice while leaving the moderator in charge of the process. Widely practiced in marketing research and used as a qualitative method for generating data in the social sciences, such methods elicit information that is biased by the management of the group’s interactions, revealing something very different from what people would express in unconstrained conversations.

**Inauthentic questions**
I suggested that asking questions with unknown answers creates possibilities for participants to choose their contributions and experience respect when their answers are acknowledged by responding to them. But questions may be inauthentic as well. Knowledge tests, for example, whether administered in educational settings, aptitude tests for hiring employees, or scientific research, are not geared to understanding but to establishing a respondent’s comprehension, the criteria for which reside in the questioner. Asking questions to which the answers are known is consistent with conceptualizing communication as the accurate transmission of information from one mind to another – a process that is institutionalized in many educational and administrative situations that have nothing to do with conversation.

In public opinion research, interviewees are asked to commit themselves to answer an interviewer’s questions, and to give up their conversationally expected ability to ask questions of their own. In this genre of social research, questions are standardized for all interviewees, asked according to a schedule, and a prepared set of answers conforms to the interest to the sponsors of the research. Whatever results from such interviews has less to do with what people would say to each other than with what sponsors want to hear (Krippendorff 2005) – a seriously biased investigative technique. Talk show hosts on radio or television are notoriously in charge of what counts as appropriate to the institutionalized genre they enact. They define the topic, ask the questions, interrupt as they see fit, including signaling the audience to applaud. Talk show guests tend to go along with these inauthenticities for the publicity this affords them on a show.

**Institutionalized interactions**
Mariaelena Bartesaghi (2009a), studying therapists’ use of questions during therapeutic sessions, found less obvious inauthenticities. The therapeutic use of questions may give clients the impression that the therapist is genuinely interested in their problems, but systematically directs the clients’ answers to where therapists want to go with them. She defines therapy as an institutionalized form of interaction. Therapy includes avoiding answering clients’ questions, for instance: Client: “Why can’t I see you on Monday?” Therapist: “That seems to disturb you, doesn’t it?” (Lakoff 1990: 69).

**Referring to participants in terms of stereotypical categories**
When addressing each other or some participants in social categories, for example, as a (typical) woman, black, Frenchman, gay, mental patient, catholic, or consumer, the ensuing interaction is no longer among mutually respecting individuals but between social categories in terms of which participants are expected to reply. Peter Berger and Thomas Luckmann (1966) discuss these social categories as “typpifications.” It would be difficult to hold categories accountable for what their members say and do. Similarly, when participants in conversations come to divide themselves into opposing camps with ideological, party political, or ethnic labels, for example, into progressive and conservative politicians, often resulting in the use of plural pronouns – the inclusive “we” and the exclusive “they” – communication becomes interactions among publically identified collectivities and conversation is, at best, a wrong metaphor. Party politics attests to perfectly reasonable individuals adopting ideological voices.

Even deliberately avoiding public stereotypes can degrade authentic conversation. John Jackson (2008) explores the unintended consequences of political correctness in the United States. By confining the use of racial stereotypes to conversations in the privacy of
one’s home, public discourse becomes disingenuous and the realities it constructs schizophrenic, not resolving the racial tension that political correctness was thought to alleviate. This phenomenon also exemplifies how the invocation of normative theories about proper talk in public can destroy the authenticity of conversation.

**Institutionalizing reality**

Bartesaghi (2009b) identifies several strategies that therapists apply to establish their authority vis-à-vis their clients. Some authority is already presupposed in the very act of clients seeking therapeutic advice. But in therapy sessions, this authority needs to be realized in talk. Therapeutic authority derives largely from using a vocabulary that is institutionalized in therapeutic discourse in which therapists claim expertise. Therapists are trained to reframe clients’ personal narratives in professional terms, constructing a psychotherapeutic reality for them that therapists can treat with the institutional resources they command and clients lack. This practice renders clients as incompetent narrators of their own world. Bartesaghi made three important observations. (1) The therapists she observed managed to prevent being held accountable to their clients by hiding behind the professional community of therapists, referring to themselves in terms of the collective “we,” having “years of experience,” and professional affiliations. That community is physically absent from the therapeutic session. Channeled into the conversation by the therapist leaves the client no chance to address that community directly. (2) By applying institutionally established therapeutic theories to the social life of clients – theories of the clients’ mental and emotional states that they are not expected to know – client accounts are rendered unreliable or flawed. This gives therapists (3) the justification for replacing clients’ narratives, feelings, and social problems with institutional accounts that enable treating clients as individuals by therapeutic means.

Therapeutic discourse is not the only discourse that constructs institutional realities that clients are asked to accept on the therapists’ authority and with their help. Scientists, too, tend to claim possession of the instruments for establishing objective truths, realities that laypersons must accept on account of the scientific authority articulating their truths. Teachers assume their authority vis-à-vis their students by claiming to have valuable knowledge that students need to acquire. Literary scholars presume the ability to interpret texts in ways that untrained readers cannot and authors may not be aware of. For example, Paul Ricoeur’s (1970) “hermeneutic of suspicion” insists on characterizing authors as hiding their agenda behind their writing, which has given literary scholars the professional license to construct what could underlie a text regardless of what its author says it means. In effect, this scholarship thrives on institutionalizing what has been called “conspiracy theory.” It permits scholars to not listen to how others – readers and authors – interpret the text they are analyzing. Conspirators must, by definition, deny being one. It follows that an author’s denial of the suspected intentions can be interpreted as evidence for the validity of the suspicion – a cognitive trap. One cannot converse with institutionalized realities, only with people willing to consider them as mere hypotheses, which is what social constructivism advocates.

Not confining accountability to those present in conversations is a premise of sociological theorizing. Besides what I mentioned above, there are at least three ways this can happen and it would be important to recognize the linguistic ground, as Habermas (2001) does, that makes sociology possible.

**Speaking for absent others**

When therapists rearticulate their client’s stories in therapeutic terms, therapists and clients are at least co-present. It is conceivable; therefore, that they could hold each other accountable should the evolving conversation go astray. Even institutionalized realities can be contested, although I am told that clients in therapy rarely ever do this in their sessions, which is not to rule out the possibility of expressing their misgivings in conversations with trusted friends. However, when speaking for absent others, speakers usurp the voices of individuals who, perhaps conveniently excluded from a conversation, can neither be questioned within that conversation nor be held accountable for their views as channeled into a conversation by one participant. Noble intents notwithstanding, speaking for the poor, oppressed, minorities, victims of crime, or even for familiar acquaintances is a discourse strategy in which speakers claim to have more voices than their own. When compellingly asserted, this gives speakers rhetorical strengths over those who cannot claim such backing. Reporting rumors or something overheard may not have much weight, but claiming to speak for one’s boss during a contentious meeting can convert a conversation among equals into a game of usurped, claimed, perhaps invented voices, which is no longer between authentic participants.

**Speaking as representatives of others**

I.e., individuals, organizations, movements, or governments. Lawyers represent their clients in court mainly because untrained individuals believe they do not have the know-how to navigate themselves through the legal system. In taking on a case, lawyers translate their client’s stories into legally valid narratives that a court is designed to handle and to which clients are asked to submit for fear of failing. In this process, clients become legal categories – plaintiffs, defendants, or witnesses – whose roles are circumscribed by being treated as their category and forced to respond accordingly. Politicians in democratic governments often face the difficult choice between speaking their conscience or in the name of the constituencies that elected them. The latter has the advantage of giving those with larger constituencies more clout and allows them to defer voting until after consulting with their constituency. In all of these cases, interactions are constrained by the process of representation. Therefore, a parliament is not a place for conversations but for institutionalized debates, public posturing, behind door negotiations, compromises, and voting in the name of absent others.

**Speaking as the occupant of an office**

In social organizations, members are assigned to offices that serve particular functions with responsibilities for coordinating the work of subordinates. Occupants of an office dedicate all communications to the purpose of that office, speak from that position, not for themselves, and expect all subordinates to be accountable to them without challenging their office. The transitivity of such asymmetrical accounting practices cre-
While these discourses can be enormously productive, I do suggest that conversations open spaces for people to realize each other as human beings, that conversational competencies precede discursive practices developmentally (children need to acquire conversational competencies before becoming competent in a particular discourse), etymologically (the vocabularies of discourses tend to go back to generations of speakers), and epistemologically (personal experiences that enter conversations may become displaced by discursive constructions of reality). Therefore conversation should not be ignored when theorizing human communication in general and human participation in social organizations (Krippendorff 2008b), science, and culture, in particular.

According to earlier distinctions, there are five constitutive features of discourse.

- **Discourses surface in the artifacts they construct**, including the body of their texts. The discourse of physics constructs a logically consistent universe amenable to observation and causal explanations; that of medicine, one of diseased or debilitated human bodies open to cures or surgical interventions; that of design, one of future technologies of everyday life. Discourse-specific vocabularies are standardized building blocks for constructing such artifacts. The body of artifacts that a discourse attends to needs to remain open to rearticulation, recombination, and creative extensions, or else the discourse dies for lack of space. Traditional discourse analysts limit their attention to available texts. I maintain this to be insufficient. Texts are read and embedded in talk among particular people and acted upon. The artifacts that discourses generate include all of their visible and somewhat enduring manifestations, not just texts but also discourse-specific universes, professional practices, and technologies. These artifacts are co-constructed in interpersonal interactions, which, while inconceivable without individual cognition are not intelligible in terms of cognitive processes.

However, unlike the traditional emphasis of discourse analysis, these artifacts alone are not sufficient for understanding the operation of a discourse; hence there are four additional features of discourse.

- **Discourses are kept alive within a community of their practitioners**. Texts need to be read, reread, reinterpreted, reconstructed, and updated by members of a discourse community and maintained by members of a discourse community of its practitioners.

Discourse as constrained conversation

I use "discourse" to describe what conversations can become when untoward conversations move are not accounted for or repaired. Discourse surfaces when interactions become systematized, organized, institutionalized, and no longer open to everything its participants may have to say; when dialogical equality is replaced by asymmetrical communications; when the insistence on consistencies constrains the creativity of authentic conversations afford their participants; and when self-organization (communicational closure) is replaced by hierarchies of asymmetrical accounting practices outside the present interactions. Elsewhere, I have written about “discourse as systematically constrained conversation” (Krippendorff 2009: 217–236) of which I can outline here only its principal features.

To be clear, when saying that conversation descends, degenerates, or erodes into discourse, I do not wish to imply that discourse is an undesirable form of language. We know many discourses that have made contemporary society more livable. We have reasons to be proud of scientific discourse, public discourse, legal discourse, design discourse, and the discourse of cybernetics (Krippendorff 2008c), to name but a few. While these discourses can be enormously productive, I do suggest that conversations open spaces for people to realize each other as human beings, that conversational competencies precede discursive practices developmentally (children need to acquire conversational competencies before becoming competent in a particular discourse), etymologically (the vocabularies of discourses tend to go back to generations of speakers), and epistemologically (personal experiences that enter conversations may become displaced by discursive constructions of reality). Therefore conversation should not be ignored when theorizing human communication in general and human participation in social organizations (Krippendorff 2008b), science, and culture, in particular.

According to earlier distinctions, there are five constitutive features of discourse.

- **Discourses surface in the artifacts they construct**, including the body of their texts. The discourse of physics constructs a logically consistent universe amenable to observation and causal explanations; that of medicine, one of diseased or debilitated human bodies open to cures or surgical interventions; that of design, one of future technologies of everyday life. Discourse-specific vocabularies are standardized building blocks for constructing such artifacts. The body of artifacts that a discourse attends to needs to remain open to rearticulation, recombination, and creative extensions, or else the discourse dies for lack of space. Traditional discourse analysts limit their attention to available texts. I maintain this to be insufficient. Texts are read and embedded in talk among particular people and acted upon. The artifacts that discourses generate include all of their visible and somewhat enduring manifestations, not just texts but also discourse-specific universes, professional practices, and technologies. These artifacts are co-constructed in interpersonal interactions, which, while inconceivable without individual cognition are not intelligible in terms of cognitive processes.

However, unlike the traditional emphasis of discourse analysis, these artifacts alone are not sufficient for understanding the operation of a discourse; hence there are four additional features of discourse.

- **Discourses are kept alive within a community of their practitioners**. Texts need to be read, reread, reinterpreted, reconstructed, and updated by members of a discourse community and maintained by members of a discourse community of its practitioners.

---

**The Author**

Klaus Krippendorff is the Gregory Bateson Professor for Cybernetics, Language, and Culture at the University of Pennsylvania’s Annenberg School for Communication. He holds a Ph.D. in communication from the University of Illinois, where he studied with W. Ross Ashby, and a graduate degree in design from the avant-garde, now defunct, Hochschule für Gestaltung, Ulm. He is a Fellow of the American Association for the Advancement of Science, the International Communication Association, the East-West Center in Hawaii, and Netherlands Institute for Advanced Studies. He is a Past President of the International Communication Association, founder of the International Federation of Communication Associations, and active in the American Society for Cybernetics. Krippendorff has published over 100 journal articles and several books on communication, social science methodology, system theory, cybernetics, and design. Among his books are The Analysis of Communication Content (co-ed.); Information Theory; Content Analysis (translated into everal languages); Communication and Control in Society (ed.); A Dictionary of Cybernetics; Design in the Age of Information (ed.); The Semantic Turn: A New Foundation for Design; The Content Analysis Reader (co-ed.); and On Communicating, Otherness, Meaning, and Information. He is currently exploring the role of language in the social construction of selves, others, and social organizations; issues of conceptual entrapment and emancipation; the human use of cyberspace; and the design of future technologies.
munity specializing in that practice. Texts have no meaning without readers and the artifacts of a discourse are rendered meaningful primarily by the members of a discourse community that has created and used them in their midst as well as by users outside the discourse. A discourse community is self-organizing by legitimizing its own practices, including creating and maintaining standards for reading, writing, interpretation, and construction of their own realities, conditions for membership in the discourse community, and criteria for attributing meanings to the activities of its members. For example, the medical discourse community trains future members, certifies its practitioners, determines codes of conduct and defines the criteria for good medical research. All discourse communities are autonomous and pursue their distinct identities.

- Discourses institute their recurrent practices. This is to say that discourse-specific practices – courses of education, applicable methods and techniques, media of publications, awards for outstanding accomplishments, etc. – are codified, institutionalized, and maintained as the preferred practices of members of the discourse community and maintained in the name of that community. Social science publications, for example, are carefully evaluated by editors and reviewers, encourage a common vocabulary, allow younger members to qualify for promotion, and assure the efficiency of constructing discursive artifacts. Theorists refer to their predecessors, research methods build on each other, intervention strategies are improved over time – all of which contribute to an institutionalized history of discourse practices, which has the benefit of avoiding the duplication of innovations, standardizes methods, typifies expertise, and thus serves to make the discourse more efficient.

- Discourses draw their own boundaries, deciding who and what belongs and what does not. Some discourses identify themselves by reference to the construction of a particular class of artifacts – biologists, for example, are concerned with what they construct as living organisms; others are committed to applying particular theories – physicists, for example, are committed to causal explanations and the construction of a consistent universe; still others are dedicated to solve particular problems – engineering, for example, seeking technological solutions to all kinds of problems, including social ones.

- Discourses justify their practices to outsiders. Justifications may be motivated by the need to continually recruit new members for the discourse community to remain viable, mobilize the resources necessary to construct their artifacts and promote their use by others. But justifications also provide the perhaps unintended ground for driving various discourse dynamics. One may note discourses that compete with one another, as science and religion did until the discourse of religion found a niche that resists scientific penetration. Some discourses consider themselves foundationalist, such as physics claiming that everything real is physical in nature and everything else is inferior science or fiction. Some discourses colonize others, as cognitive science has been doing lately to psychology.

**Computation**

If discourse emerges when constraints on authentic conversation are naturalized, talk becomes institutionalized, and unequal accounting practices are accepted and directed to the construction of discursive artifacts, then the implementation of technological solutions to social problems or the replacement of social practices by more efficient mechanisms can be considered a move from discourse to the entirely non-linguistic processes of computation. Today, we are witnessing the massive translation of discursive practices into efficient computational mechanisms: delegating repetitive work to robotic devices, searching for relevant texts on the internet with search engines, scheduling airline traffic, letting computers buy and sell stocks, using online accounting for the essential variables of social organizations, and automating whole businesses. In the same way, statistical software in the social sciences has replaced seemingly endless and error prone hand calculations by teams of researchers, and electronic banking accomplishes what a social network of coordinated bank employees did before the advent of computers. These replacements are driven by the increasing availability of software, discursively developed by armies of collaborating programmers.

Software is written in a computer language and explicates algorithms, i.e., step-by-step instructions in which all conceivable paths are anticipated and by means of which receptive hardware can be programmed to be a purposefully functioning machine. Much like in discourse, where it does not matter who practices it as long as someone does, computation is not tied to particular material manifestations as long as it works. In other words, the material makeup of hardware is irrelevant to its proceeding from state to state in a determinist fashion. Hence, software specifies a deterministic process, rendering computers deterministic machines that cannot choose what they do. They have no agency. Non-digital technologies – simple tools, cars, hospitals, public performances – may not be programmable as computers are, but their design has always focused on how they go from here to there, what, in the digital world, is called “computation,” hence my use of this term.

All conversations, discourses, and computations produce something. The products of conversations and discourses are still coordinated by talk, text, and interactions. Computations, however, once initiated by human actors, run their course unless intervened with at their interfaces. People may blindly accept the results of computations and allow themselves to be affected by these devices, but this is a user’s choice, not a necessity.

Because of the difficulty of grasping the complexities of computational devices, we often attribute human qualities to them – intelligence, temperaments, likes and dislikes (Turkle 1984, 2007; Reeves & Nass 1996), and the ability to act (Latour 2005). However, such attributions do not change the deterministic nature of computational artifacts. One cannot hold computers accountable for what they do. Therefore, replacing discursive practices with computational technologies and relying on them in everyday life amounts to a fundamental shift away from human participation. It is truly amazing to realize how many discourses depend on digitally mediated communication and computation and how little the social sciences have conceptualized this fact of social life or how they have confused the two as Latour (2005) does. Here, cybernetics has much to explore and many insights to offer.
Conclusion

To sum up, Figure 1 depicts a continuum between the extremes of authentic conversation and computation, populated by discourse formations of varying degrees of rigor. Conversational competencies include, as I suggested, the ability to repair untoward moves that speakers may make, which can bring discourses back to conversations and the latter to authentic ones. But by not repairing problematic encounters when they occur, by consenting to limit accountability for problematic actions, conversations unwittingly drift into discursive forms that may well construct realities of a kind that conversations cannot construct – think of sophisticated information systems, highways, and the infrastructure of cities. The evolution of such artifacts is accomplished by discourses that coordinate large numbers of human participants, including over some time. It follows that social artifacts of such complexity cannot possibly be explained by the cognitive constructions of an observer or of any one of its constituent creators, users, or stakeholders. What participants do know is their own creative but always only partial contributions. The remainder consists of trust in the linguistic competence of the other participants to coordinate their understanding and interact towards what is to be done. In the transition from conversation to discourse, conversational possibilities are traded for practical conveniences. In the transition from discourse to computation, seemingly costly, unpleasant, or inefficient discursive practices are implemented in mechanisms whose ultimate consequences may be difficult to foresee.

I am suggesting that the move from conversation through the large domain of discursive forms is attracted by the ultimate temptation of turning social processes into productive algorithms whose operation in various technologies is no longer social, except before their inception and subsequently, at occasional interventions through multi-user interfaces with them. Since computational artifacts often are beyond individual understanding of how they work, such technologies can no longer be treated as tools under rational control of their creators and users. Uncritical reliance on computation can lead communities into unintended realities that may well become unbearable to live in and therefore constitute an important domain of scholarly and designerly attention.

This essay is intended to expand the limits of radical cognitive constructivism, which confines itself to individual understanding, into the social domain, and introduce doubts in the epistemological position of observers at the expense of participatory and interactive reality constructions. I maintain that human realities, including the idea of cognition, are conversational or discursive realities in the sense that we humans interactively participate in their construction – without being in charge or fully cognizant of each other’s conceptions, except for our contribution to them.

I hope that readers of this essay consider conversation – not individual cognition and efforts to describe one’s observations – as the essentially human way of living together. Following from that, is the awareness of the often casually accepted drift from conversation through various discursive forms to computation. I invite readers to draw finer distinctions within the domain of discourses and reflect on how their own contributions affect the spaces left to exercise accountability along this sometimes appealing journey. Although computation deserves more attention than I could devote here, it should be recognizable as that which early cybernetics thrived on and proposed in the form of theories, models, and mechanisms for augmenting social reality. Computation undoubtedly has vastly expanded the horizon of our abilities, but it can also constrain human agency. When moving through various discourses, converting recurrent social practices into computational artifacts, we should always preserve the possibility of returning to authentic conversation, its sheer pleasure and fundamental humanness.

Acknowledgement

I am grateful for constructive comments by two reviewers, Mariaelena Bartesaghi, and an unidentified one.
References

Nofsinger R. E. (1999) Everyday conversa-

Received: 29 April 2009
Accepted: 26 June 2009
Conversation, Individuals and Concepts
Some Key Concepts in Gordon Pask’s Interaction of Actors and Conversation Theories

Bernard Scott  Cranfield University, UK <b.c.e.scott@cranfield.ac.uk>

**Purpose** – Gordon Pask has left behind a voluminous scientific oeuvre in which he frequently uses technical language and a detail of argument that makes his work difficult to access except by the most dedicated of students. His ideas have also evolved over a long period. This paper provides introductions to three of Pask’s key concepts: “conversation,” “individuals,” and “concepts.”

**Method** – Based on the author’s close knowledge of Pask’s work, as his collaborator for ten years and as someone who has had access to and is familiar with almost all of his published work, the paper selects three of Pask’s key concepts for elucidation in order to motivate the interested reader to explore Pask’s work in the original.

**Results** – Pask’s key concepts, “conversations,” “individuals,” and “concepts,” which are central in his conversation theory and his later elaborations in “interaction of actors theory,” are shown to be grounded in fundamental principles from cybernetics. Furthermore, the form of Pask’s theorising is that of second order (reflexive) meta-theorising, developing theories of theorising that explain their own form and genesis.

**Key words** – Conversation theory, conceptualisation, individuality, interaction of actors.

**Introduction**

Gordon Pask’s theorising evolved over several decades. His first academic paper (Pask 1959) already sowed the seeds of his major concern: what concepts are and how they are acquired, addressing both questions of embodiment of the processes in material substrates, such as a brain, and the form of the process as cognition or “mentation,” a term frequently used by Pask in those early years. The full outline of his concerns is set out in a major paper (Pask 1965) as a holistic “global cosmology of consciousness” (our term). As described elsewhere (Scott 2007b), his articulation of “conversation theory” (CT) as a formal account of his theorising to that date appeared in the early 1970s. (For a definitive statement, see, Pask 1975a). The later elaborations, which he referred to as “interaction of actors theory” (IA), were largely made in the late 1980s and early 1990s. His last publication (Pask 1996) was a terse, formal summary of the main tenets of CT/IA.

In this paper we summarise what we see as some of his key concepts that, through various reformulations and changes in terminology, remained central in his thinking. The concepts focussed on are those of “conversation,” “concept” and “individual.” All three are central in CT and his later elaborations in IA. We show how his key concepts are grounded in fundamental principles from cybernetics. We note that the form of Pask’s theorising is that of second order (reflexive) meta-theorising, developing theories of theorising that explain their own form and genesis.

First we say a little about cybernetics and Pask’s embracing of it as one who consistently referred to himself as a cybernetician. We then review Pask’s early work on adaptive teaching to show how his concern with human learning evolved into a general interest in human cognition and communication and the formulation of a general theory – conversation theory – which is briefly overviewed as a second order theory, i.e., as a theory of theory building.

We then look more closely at the emergence of selves and others and the conversational dynamics thereof. We discuss Pask’s concepts of psychological and mechanical individuality as a precursor to overviewing what Pask refers to as “conceptualisation” and his later formulations of CT as IA. Finally, we briefly examine the limits and potentials of knowing and being.

**Gordon Pask and cybernetics**

Pask was internationally known as a cybernetician and referred to himself as such. He embraced cybernetics both as a transdisciplinary field, bringing order to a variety of disciplines by revealing similarities and differences, and as a metadiscipline, a discipline about disciplines that gives an account of how disciplines come into being. He is known as one of the first to promulgate cybernetics in the UK (see Pask 1961). His commitment and contributions to cybernetics are well documented in Scott (2007b, 2008), Glanville (1993), and Scott & Glanville (2001a, 2001b). Indeed, amongst his peers in cybernetics he was known as the cybernetician’s cybernetician (Beer 1993; Foerster 1993). Beer quotes, with approval, a Pask aphorism about the activities of cyberneticians: “It is the science of defensible metaphors” (Beer 1993: 16).

In his understanding of cybernetics, Pask (1961) accepted various formulations as compatible ways of defining the field. These include the Macy Conferences’ emphasis on “Circular causal and feedback mechanisms in biological and social systems,” Wiener’s (1948) “Control and communication in the animal and the machine” and Ashby’s (1960) “The study of all possible machines.” Pask also readily embraced von Foerster’s (1974) distinction between levels of inquiry within cybernetics: first order study of observed systems and second order study of observing systems. He understood the two kinds of study as complementary and saw the distinction between them as already imminent in the
thinking that pervaded the Macy Conferences.

Pask agreed with von Foerster’s aphorism that, “The goal of second order cybernetics is to explain the observer to him/herself” (Foerster 1974). In developing CT and its precursor theorising, Pask had already accepted this reflexive aspect of cybernetics and had addressed the epistemological issues that arise when taking this reflexivity seriously. Pask (1969) sets out the research agenda that led to the formulation of CT.

Early work on adaptive teaching

At the time when teaching machines and programmed learning first became fashionable, following the writings of the behaviourist, B. F. Skinner, Pask articulated the cybernetic view that “Teaching is the control of learning” and designed and built a number of machines that did, indeed, teach in the required sense. At the same time, his systems served as an, what Pask called, “embodiment” for the theories of learning and conceptualisation that he was evolving.

The systems in question have a long history, which has been well documented by Pask (1975a) himself. Their development excited international interest and parallel work was carried out by, among others, Lawrence Stolurow, Brian Gaines, Edwin Hudson, and Charles Kelley. As Gaines notes in his survey of the field, “Pask has made available a very deep and comprehensive discussion of automated training and has placed it in the general context of interactions between self-organising systems” (Gaines 1967: 7).

The essential feature of an adaptive teaching machine is that it monitors a learner’s performance and poses problems at the level of difficulty most conducive to effective learning: as the student’s performance improves, so the level of difficulty is increased. This may be achieved in several, parallel ways by several feedback loops (for example frequency of stimulus presentation, complexity of stimuli presented). Where the skill has a clear hierarchical structure, higher order feedback can monitor and integrate distinct subskills. Pask and Lewis (1968) argue that such adaptive, responsive environments are what are required to observe a self-organising system, referring to the methodology as a “null-point” or “steady-state” approach. The admission of a flaw in the methodology led to the first systems in which the full logic of a conversational interaction was embodied, albeit in embryonic form.

The basic assumption of the earlier systems was that the student would abide by the agreement to attempt to perform the skill effectively at all times. Pask and Lewis noted a tendency for some subjects to override the adaptive controller by selectively making deliberate errors. In this way, the subject herself could determine which component of a skill was rehearsed. At first this was seen as a form of play. However, further analysis and interviews with subjects provided evidence that such subjects were imposing their own learning strategies on the system. A series of studies (Pask & Scott 1971, 1972) investigated the form of learning strategy adopted in a free-learning situation, where the student was free to focus her attention on different components of a structured skill. Two classes of strategy were detected: a holistic, many-at-once approach, and a serialist, one-step-at-a-time approach, either of which could be effective.

A “conversational” teaching system was devised wherein the student’s choices were accepted only if her performance was at an acceptable level. If not, the teaching system took control, basing its decisions on empirically validated diagnostic measures. These “conversational” systems were shown to be more effective than either free-learning alone or straightforward adaptive machine teaching.

These studies were the inspiration for a long series of investigations of individual differences. At the same time, they helped inspire Pask to elaborate his theory of learning and teaching so as to fully allow for the student to act as her own teacher within her own descriptive framework for the structure of a skill or body of knowledge.

Conversation theory

The basic conversation theory model is shown in Figure 1: the “skeleton of a conversation.” It shows a snapshot view of two participants (in this case, learner and teacher) in conversation about a topic. The model distinguishes verbal, “provocative” interaction
Individuality

In CT, a distinction is made between psychological individuals (“p-individuals”) and mechanical individuals (“m-individuals”).

P-individuals are conversational processes but may themselves be participants in a conversation. In contrast, a mechanical individual is a system that is observed as an “it.” Notice it is we as external observers who distinguish the two sorts of system. With respect to human beings, we commonly shift perspectives from one form of individual to the other. A surgeon may talk to Joe Bloggs, the p-individual, before and after his operation, but during the operation, he is interacting with an m-individual, an “it.” Pask often remarked that his chief interest in m-individuals was their potential for embodying and supporting the processes of p-individuation.

A typical m-individual is a processor, such as a brain/body system, and its augmentations, such as a notepad or personal computer. Such a processor may execute one or more p-individuals as processes. Here, Pask has in mind the internal conversation that is cognition, most clearly observed in the experimental free-learning situations where the student, acting as her own teacher, directs her own attention. Alternatively, several m-individuals may execute one p-individual, as in the conversation that constitutes a social organisation.

It is important to stress that m-individuals do not have “beliefs.” Holding beliefs is a property of p-individuals, which may rightly be referred to as “belief systems,” synonymous with a “conceptual system.”

The great advantage of making and maintaining the distinction between p-individuals and m-individuals is that the observer is not obliged to maintain a one-to-one correspondence between them. An m-individual may embody several p-individuals; several m-individuals may embody one p-individual. As noted, the distinction between the two types of system is one that is made by an external observer. However, they are both in their own ways self-distinguishing by dint of organisational closure. As detailed below, a p-individual is a self-distinguishing, self-reproducing, organisationally closed collective of concepts. M-individuals, as typified by biological systems, are also organisationally closed, self-distinguishing and self-producing. They are the autopoietic systems of Humberto Maturana (Maturana & Varela 1980).

At first sight, the difference between p-individuals and m-individuals may look like some form of Cartesian dualism of mind and body. But this is not so. Pask believed neither in disembodied minds nor disembodied bodies. With Warren McCulloch and the Stoics, he would agree that, “A thought in the head is like a fist in the hand” (McCulloch 1967: 349).

CT brings unity to the discipline of psychology by unifying process, individual and social psychologies. CT also provides a unified view of psychology, sociology, anthropology as “sociocybernetics.”

Within psychology, CT brings understanding and reconciliation to two major approaches: the study of persons and the study of processes. There are some psychologists that emphasise that it is important to study the whole person as an integral whole; sometimes these are known as humanistic or individual approaches within psychology. Approaches that tend to emphasise the whole person include those of Maslow, Kelly, Freud, Jung and Adler. In contrast, the process approach is one that studies particular aspects of the whole person as separate domains of study. Examples are studies of perception, of problem solving, of memory, and of language processing. Within these domains of study there are further subdivisions, for example, auditory perception, visual perception, and olfactory perception. Pask brings order to the person/process distinction using the classic cybernetic technique of abstracting over sets of special cases. In the process approach, what is being studied are particular psychological processes. What they have in common is they are, in fact, all processes. They can all be understood or modelled as procedures that select and act on “capta” and carry out operations to transform them in some way. An image is recognised, an event is remembered, a problem is solved. Pask’s general name for all these different sorts of psychological process is “a concept.” As discussed further in the next section, a p-individual is a system of processes that are “organisationally closed,” so that they are not only productive but are also incidentally but necessarily self-productive, that act so as to reproduce themselves as a coherent whole.

What is also innovative and unifying in CT is the voiding of the distinction between the human individual and the social processes that are constitutive of him/her and which he/she constitutes. Margaret Mead (1934) posited that “the self is a social process.” Pask agrees with Mead, Vygotsky, Buber and von Foerster that the psychological individual is dialogical in form, is a social process, is constituted by an inner dialogue, is an inner conversation. As a good cybernetician, Pask abstracts from specific cases and voids the dis-
tinitions and thus argues that all conversations, all dialogues, all social processes are psychological individuals. They are all organisationally closed, self-producing, collectives of concepts. 3

Conceptualisation

There is a general view from the disciplines of anthropology, psychology, education and cognitive science that “to have a concept” means one has the ability to assign a particular stimulus item to a particular class or “conceptual category.” It is also usual to make a distinction between the cognitive abilities of animals and humans. Whereas it is acknowledged that animals, like humans, can discriminate between and generalise over particular stimulus items and thus assign them to particular “stimulus classes,” humans also have the ability to name such classes using an appropriate “verbal label” and may also be directly taught, by verbal instruction, what the rules are that govern class membership.

Although some authorities refer to animals as “having a concept,” here, in line with Pask’s theorising, our attention is focussed on human conceptualisation, where concepts may be tacit or may be named and where there may also be explicit rules for class membership. Acquiring a new concept is sometimes referred to as “concept attainment.” New concepts may be acquired by having the rules for membership made explicit. Alternatively, new concepts may be acquired by inducting a rule from the presentation of exemplars. This was the procedure followed by Bruner, Goodnow & Austin (1956) in their classic studies of concept attainment. A set of stimulus items is presented and the experimental subject has to identify those that belong to the experimenter-defined conceptual category. The subject is provided with feedback, “knowledge of results,” as to whether or not she is correct in assigning a stimulus item to a particular class. Following this procedure, the subject becomes successful at making the correct discriminations and may or may not be able, eventually, to state the rules governing category membership. Being able to classify items without being able to state the rules is sometimes referred to as having a “tacit concept.”

Following the work of Bruner, Goodnow and Austin and the work of behaviourists experimenting with animals, a large literature grew up on how to teach concepts (see, e.g., Engelmann 1969; Klausmeier, Ghatala & Frayer 1972).

Drawing on the early work of Whorf, Sapir and others on the relationship between language and conceptual categories, anthropologists have also contributed to the literature on conceptual categories, particularly with respect to the question of whether or not there are “universals” in human conceptualisation or whether concepts are culturally relative (see, e.g., the work of Rochs 1973).

In the 1970s and onwards, with the advent of work on artificial intelligence (AI), the field of “cognitive science” emerged. AI researchers looked to build “expert systems” capable of solving problems, including problems of classification and the induction of rules. Psychologists used computer programs as a way of building models to simulate human conceptualisation and problem solving.

Two broad classes of model have emerged, so-called “symbolic” models and “sub-symbolic” (also known as “connectionist”) models. The former typically use computer languages such as LISP or Prolog to represent concept rules. The latter typically simulate brain activity by training networks of idealised neurons to discriminate and generalise across stimulus items. Some researchers use hybrid models with both “symbolic” and “sub-symbolic” aspects. In “symbolic” models, concept rules are stored and retrieved from “memory.” In “sub-symbolic” models, concept rules are stored (learned and remembered) as changes to the forms of connectivity between neurons in a network, such as the “weightings” that determine whether or not a particular input will induce a neuron to fire, and thus transfer an input to other neurons. The network is trained to recognise patterns (form concepts) by being shown exemplars as input for which it must give an appropriate output. Correct responses are “rewarded” by having the weightings changed so that on future occasions the network will tend to respond in a similar manner.

Despite these massive research efforts, the question of where new concepts come from is not satisfactorily answered by the “transmission of culture” models of anthropologists and educators nor by the “acquisition and application of rules” models found in mainstream cognitive science. In philosophy and logic, the creation of a new conceptual category is sometimes referred to as a new “predication” or as “abduction” (following Charles S. Peirce) or the “making of a distinction” (Spencer-Brown 1969). In psychology, reference is made to “eduction” (following Charles Spearman). As we shall see, this question of the formation of new concepts is central in CT.

In CT, a concept is defined as a procedure that brings about, maintains or recognises a relation. A concept may be likened to a program or operator that solves particular problems. “Relation” is used here as an empty slot or label for that which is being acted upon by the process as input or product. Recursively, there are concepts whose domain of application, whose input and products, are other concepts. This affords the construction of hierarchies of concepts. Thus, there can be problem solver concepts, the task of which is to construct and select from amongst lower-level putative problem solvers as an evolutionary process guided by feedback from the problem domain about the success or not of their application. This is one of Pask’s very elegant definitions of what learning is: the construction of a hierarchy of problem solvers (Pask 1975b). 4

A system of concepts may also be self-producing. Instead of a hierarchy of processes, we now have a heterarchical system of processes that are organisationally closed, that not only are productive but are also incidentally but necessarily self-productive, that act so as to reproduce themselves as a coherent whole. As noted above, Pask refers to such totalities as psychological (p-) individuals and distinguishes them from the mechanical (m-) individuals (biological systems and their augmentations, possibly other kinds of system) that are the processors that embody the p-individuals. This is the dynamic of p-individual conceptualisation. This dynamic is distinguished as different from the dynamic of m-individual adaptation. In m-individuals there is, internal to the system, a variety generation of the “conversation” between “parts” of a self-organising (pace, organisationally closed) system. This latter “conversation” is an analogue of the conversation that is the p-individual but is not, unlike the latter, something that, as observers, we can participate in. (We can of course affect the m-individual by various forms of intervention such as surgery or behaviour shaping techniques).
Unlike CT, most other psychologies do not make the p-/m-individual distinction. The closest is the analogy “thought is to brain as program is to computer.” Note, none of these latter posited entities has the property of organisational closure; none is an “individual.”

CT is a reflexive metatheory, a theory of theory building that explains its own genesis. Those that do recognise the importance of individuality usually posit an identity or one-to-one isomorphism between the m- and a p- (as in various positions regarding ontological status of “minds” and “brains” or “mind” and “matter.”) The CT analytic (rather than ontological) distinction opens up the idea of one:many and many:one mappings between p- and m-individuals. As noted, this is extraordinarily useful in the social sciences as a conceptual unification and also as a theory-driven basis for empirical studies.

Conceptualisation and consciousness

Conceptualisation calls forth a world of experience. It is important to appreciate that for Pask this implies a radical anthropism in the sense that each conceptualising system invokes its own universe, with its own spatial and temporal forms, including experience of its own embodiment. Conceptualisation is prior to there being a universe to conceptualise about. Conceptual systems (p-individuals), when synchronised with each other through ‘information transfer’ in the sense of Carl A. Petri (1966), bring forth shared consensual universes in which, by agreement, spatial and temporal forms are, for the moment of the encounter at least, shared. Maturan (1995), in a parallel narrative, talks of “coordinations of coordinations” that bring forth consensual domains. Pask developed ways of imaging the dynamics of the processes of conceptualisation, the processes of concept production and reproduction, which go on continuously, or as Pask puts it, “Conceptualisation is conserved” (Pask 1996). He also describes the particular activities within that dynamic that we experience as directed thought and perception. These latter are processes of which we are usually to some extent aware and may experience with ourselves or with others as conscious events (Latin consicio, “know with” as in “know with another person”; McCulloch 1965; Lewis 1967). Pask considered this, his development of a formal theory of the form of consciousness, to be perhaps his greatest achievement (see, e.g., Pask 1981).

It is important to appreciate that Pask (following McCulloch) makes a distinction between the use of terms “awareness” and “consciousness.” Awareness is a general property of any self-organising system experiencing uncertainty. As noted above, consciousness refers to “knowing together.” Coordination evokes where the tokens “I” and “Thou” become established as shared tokens for the “objects” that are the participants. In general, objects when perceived and experienced as unities are themselves tokens for that which is conserved in interaction. An infant “plays with a ball.” The object “ball” emerges as a concept for that which has been experienced as constant in the interaction. Piaget (1972) refers to this as “object permanence.” McCulloch (1965) uses the term “computation of invariants.” Von Foerster (1974) refers to the computation “of a stable reality.”

The coordinations that lead to the “I” and “Thou” referencing are thus those in which the participants compute (model) each other as self-organising systems, as systems that distinguish objects, and, at a higher level of recursion, as systems that are self-distinguishing. It is at this point, where a self-organising system distinguishes itself as a distinguisher of self and other objects, that the system can be said to be “conscious” in the sense of knowing with itself that it is a self. Up to this point, this knowing with oneself has been imminent in the “proto-consciousness” (our term) of awareness. Misheva (2009) presents an elegant argument for the case that consciousness, in the sense of having a “conscience,” emerges in triadic rather than dyadic interactions. The third participant is necessary for providing the first two participants with the awareness that their interaction is being observed. In full “self-consciousness,” this awareness of being observed is internalised.

Chapter 11 of Pask (1975b) gives an account of the dynamics of consciousness and explains why there are shifts of attention. Essentially the argument is that events about which we are uncertain make it into our awareness. As the events that we are attending to cease to be uncertain, they are dealt with automatically, non-consciously, in more procedurally proficient and efficient ways. As this happens, some novel event about which there is uncertainty will come to occupy our attention. Notice how this account adumbrates (to use another favourite Pask term) or includes within it both goal setting as a purposeful and directed behaviour and problem solution emergence as “creativity” or “intuition.”

Interaction of Actors theory

IA, as a term, entered Pask’s vocabulary at a time when he began a fruitful collaboration with Gerard de Zeeuw. Pask himself saw IA as a generalisation from CT. The chief distinction is that whereas CT is concerned with the specifics of a conversation between participants, IA addresses the broader domain of human interaction in which conversations (p-individuals) appear, disappear and reappear. An actor is an embodied p-individual.
The author’s view is that these more general concerns were always part of the understanding of what conversation is, within Pask’s research group. However, given the focus on learning and teaching of the earlier work, there was an emphasis on conversations between teacher and learner about a well-specified topic. Where it was possible to deem that, with evidence of a learner’s having come to understand the topic in question, the conversation had come to an end. Pask referred to this sort of conversation as a “strict conversation.” In distinguishing IA from CT, Pask often writes as if the earlier work was only concerned with strict conversations (see, e.g., Pask 1996). In the author’s view, this is a simplification, indeed, a distortion of earlier “conversations about conversation” within the research team.

In IA, a conversation is potentially “immortal.” Conceptualisation and conversation about selves and others does not end with a participant’s death. Within human culture they may persist forever (Pask 1991a, 1996). IA is reflexive. The theory’s author is an actor. (A Paskian actor can be considered as a more well-specified entity than an “observer” or “observing system,” both of which terms, unless further qualified, have looser connotations.)

**Conclusion**

What are the limits and potentials of Pask’s theory?

With respect to epistemological limits, Pask (1996) tersely states, “Life is ineffable and ineluctable.” This can be paraphrased as saying that the complete story of life is unsayable and being alive is un-unlockable; we cannot escape out of life to observe it from the “outside.”

With respect to potentials, Pask (1969) notes, “There is no end to forms of action and interaction.” Within these limits and with these potentials, Pask (1991) argues that we should strive for “unity without uniformity” and that, from the perspective of the cybernetician, “evil” is that which limits the right of actors to interact (Pask 1991b).

From his earliest writings, Pask’s theorising and experimental studies of human and machine learning were constructivist in orientation. This was long before the latter term became popular. Pask’s theorising is based on the concept of self-organising systems that adapt to their environments and which, through interaction with other such systems, come to know with each other. From this beginning, Pask’s developed second order (reflexive) metatheories – theories of theory building – that explain their own form and genesis.

The essence of his insights was captured in Lou Kauffman’s “A Poem for Gordon Pask” (Kauffman 2001: 726).

You kept a pattern of crumbs on your plate
For the fairies,
And imagination/infinity was
Always present in the
Crystal pool of your mind.
Through whirlwind topological tangle
Of human conversation
You found the seeds of a
Transcendent reality.
In the whisper of your voice
And the strength of your will
Description and
World described
Become
One.

**Notes**

1. For early statements about CT and learning and teaching, see Pask & Scott (1973) and Pask, actor, and Kallikourdis (1973).
2. A note about terminology: Pask, who was a friend of Maturana’s, was happy to follow Maturana’s usage and apply the term autopoiesis to biological individuals but refrained from using the term as a way of describing the organisational closure of psychological systems, preferring his own term of “p-individuation.”
3. For more on the ontogeny of a p-individual as a “social self,” see Scott (2007).
4. Already in chapter 11 of Ashby (1956), there is a discussion of the evolution of hierarchies of “regulators,” ideas similar to those of Pask. Pask’s innovation, however, is to introduce the concept of a system of regulators (qua concepts) that is organisationally closed.
5. A point to be appreciated here is that whilst CT is predicated on the “reality” of conversations taking place (i.e., it is not solipsistic), the form of any “reality” shared by participants is a matter of agreement, including agreements to disagree.
6. Scott (1987), written in 1983, takes the wider view and is prior to Pask’s use of the IA label. See also Pask (1979).
7. Many of Pask’s later papers on IA are not well-referenced within his oeuvre. However, the reader might like to consult Pask (1991a, 1991b, 1996), Pask & de Zeeuw (1992), de Zeeuw (2001), Green (2001).
8. The relationship between Pask’s theories and other constructivist approaches is addressed in Scott (2001 and 2007a).
References

Scott B. (2007a) The co-emergence of parts and wholes in psychological individua-


Consciousness as Symbolic Construction
A Semiotics of Thought after Cassirer

Hans-Herbert Kögler ♦ University of North Florida <hkoegler@unf.edu>

► Purpose – In both analytic and continental philosophy, the linguistic turn jettisoned philosophical foundationalism and gave way to a new pragmatic-hermeneutic turn regarding understanding, truth, and meaning. Yet now intentional consciousness – i.e., the relation between thought and language – still poses an issue. At stake is the convincing reconstruction of consciousness based on symbolic mediation. ► Method – In order to contribute to this discussion, the paper takes up Cassirer’s argument for the necessity of “symbolic forms” for thought. It introduces an ideal-typical notion of the sign, and develops, on this basis, the relevance of Cassirer’s arguments from symbolic synthesis and from symbolic articulation in order to finally assess the prospects of a theory of symbolic meaning after Cassirer. ► Results – The paper shows that consciousness is prone to a semiotic analysis and points the way for future desiderata of a theory of meaning. ► Implications – The results are relevant for philosophy, psychology, semiotics, as well as cultural studies. The paper projects a position of symbolic epistemic constructivism by reconstructing the mind and world, or subject-object, as emerging from the prior source of symbolic mediation. ► Key words – Cassirer, symbolic mediation, consciousness, meaning, subject/object, Peirce.

Ernst Cassirer’s Philosophy of Symbolic Forms may hold untapped promise for contemporary thought. This is because Cassirer develops a mode of symbolic constructivism that aims at overcoming the division between the mind, on the one side, and the world, on the other. According to Cassirer, the mind and the world—or subject-object—are related poles emerging from a basic process of symbolic mediation. For one, intentional consciousness is capable of understanding something as something solely by means of symbolic mediations, thus displacing the idea of a foundational consciousness. At the same time, metaphysical realism – the idea that the concept of the world as in itself structured, regardless of our understanding, makes sense—is jettisoned as untenable, i.e., reality is identifiable only with what emerges in this process. Cassirer may thus be waiting at a conceptual spot towards which both the analytical and the continental approaches in the philosophy of language are moving.

In Anglo-American thought, the early hope of a correspondence theory of truth, pairing linguistic statements and facts, has consistently given way to a pragmatic turn. From Carnap to Quine to Davidson, and from there to Rorty, Brandom, and Putnam, the linguistic mediation of reality is now seen as governed by internal, contextual, piece-meal practices of confirmation and justification. Similarly, in continental thought, the movement from phenomenology conceived as a “rigorous science” in Husserl to Heidegger’s and Merleau-Ponty’s situated understanding, to Gadamer’s historical-linguistic philosophical hermeneutics, defines a turn towards contextualism and pragmatism. In both traditions, philosophical foundationalism—either in terms of basic statements or of basic phenomena—has given way to a new understanding of the role of language for thought. But how precisely we are to conceive, according to this new pragmatic-hermeneutic turn, the intentionality of symbolic thought—i.e., the relation between thought and language—is up for grabs.

Here is where Cassirer fits in. Arguably, his work entails the most developed argument for the intentional understanding of something as intrinsically dependent on symbolic mediation. Moreover, he develops his argument from within a philosophy of mind and thus further motivates our having to take seriously language and semiotic structure when discussing consciousness. Yet Cassirer furthermore shows that we can reconstruct intentional consciousness by means of a rational philosophical analysis, thereby opposing the analytic tendency to reduce or even eliminate conscious thought in favor of a narrowly conceived naturalism or realism. This path towards a new approach of analyzing the symbolic nature of intentional consciousness is opened by a semiotic approach, by looking at how the symbolic forms of thought may allow it to exist as such. Our analysis will set the stage by introducing the idea of a semiotics of thought (1), turn then to a reconstruction of Cassirer’s core arguments in favor of symbolic mediation (2), to finally assess the viability of his approach for a comprehensive theory of symbolic meaning (3). Cassirer’s problems in ultimately arriving at a tenable philosophy of language—the most important four of which we will reconstruct in some detail—do not, however, diminish his importance for reassessing the symbolic nature of consciousness.

The idea of a “semiotics of thought”
At the core of the semiotic reflection on language is the conviction that the basic riddles of any theory of language—regarding the meaning of linguistic units, their basic identity, and the functions that allow us to categorize and analyze language—can best be understood by looking at language as made up of signs, or better, as a system of signs. Semiotics declares the nature of the sign to be central to the analysis of meaning. At the same time, it avoids the assumption that the analysis of the symbolic relation (the relation between that which stands for something and that something itself) can be analyzed in terms of a single sign or a unique relation. In contrast, semiotics proceeds from the assumption of a “holism of meaning,” suggesting that the identical meaning of a sign can only be under-
stood in the systemic context of other signs and their underlying rules (Saussure 1986).

If we are interested in the relation between thought and language, the question arises as to how the sign—that is, its inherent symbolic relation as a *quid pro quo*—relates to the intentional understanding of speakers. Focusing on the sign as such might, prima facie, suggest that we can disregard the conscious or intuitive self-understanding of agents, that we are dealing with an “objective” relation or reality here, a reality for which the “additional” (or epi-phenomenal) thoughts that speakers associate with the meaning of terms are irrelevant. We would thus be faced with the somewhat positivist claim that the sign allows us to focus on “objective” symbolic relations and contexts, and thus avoid the problem of consciousness. This would parallel the early hopes of the analytic philosophy of language, which assumed that the linguistic turn can provide us with a fully externalist account of meaning in which the intentional or subjective understanding of speakers plays no systematic role.

Regarding the philosophical approaches to semiotics and symbolic meaning, however, such a view would prove both systematically misguided and historically wrong. For one, only by drawing on one’s own understanding of what a symbolic relation is—i.e., what is *meant* when we say a sign refers to or represents something else—can the analyst begin to reconstruct and classify symbolic orders and symbolically mediated contexts. Thus, the semiotic thinker herself must make use of the intentional understanding that linguistic symbols are capable of “representing” phenomena and experiences by means of signs. And two, it is historically false that the project of a semiotic linguistics was introduced from a strictly third-personal perspective. We will provide ample proof of this by presently setting the stage for our analysis by relating it to Charles Sanders Peirce’s conception of the sign. Yet even Ferdinand de Saussure declares explicitly that we must consider the existence of a shared symbolic code, for which his version of semiotics has become as famous as it is influential, a *quasi-transcendental presupposition* for mutual communication. It is necessary if a speaker is to be able to understand a hearer. Accordingly, the intentional understanding is conceived in terms of communicative success and *as such* structurally tied to the (however implicit and unreflective) shared code. In that regard, the connection between the semiotic holism and the consciousness of the individual speaker is built into the semiotic paradigm.

What needs to be shown, to be sure, is how exactly the two spheres—the intentional orientation of a conscious speaker toward some subject-matter, implying a rational and normative dimension towards something, and the underlying semiotic holism that exists in the form of “objective” symbolic codes and rules—are related to one another. How can we understand the conscious and intentional orientation towards something as truthfully relating to this something, if this relation is essentially mediated by symbolic codes and rules? How can we do justice to the intentional claim that we talk about the world, that we reach things, states of affairs, and events by means of language, if intentional thought is not possible without a symbolic construction that mediates how we are able to talk about something, how we are able to refer?

In what follows we suggest that the semiotic philosophy of consciousness of Ernst Cassirer can help us prepare a solution to this issue. Our analysis of Cassirer’s *Philosophy of Symbolic Forms* is oriented toward its unique treatment of two assumptions that seem essential to any approach to this issue. On the one hand, higher order thought is only possible on the basis of symbolic mediation. Cassirer makes a most compelling case for how intentional consciousness requires the symbolic medium to articulate itself as thought. On the other hand, the symbolic articulation of thought can only be adequately conceived as consciousness of something, i.e., it is the- matized with regard to its function of an object- or world-relation. Just as in phenomenology, consciousness is always the consciousness of something—only now the consciousness of something is understood to be possible only on the grounds of the symbolic relation (i.e., the relation of the sign to something else). Cassirer’s project consists of what we may call a “semiotics of thought,” as he envisions solving the problem of how a symbolically mediated consciousness can refer to something in the world by reconstructing how the world itself emerges from the symbolic mediation of understanding.

Let us first introduce an intuitive conception of the sign that points in the right direction. A sign is defined, so the phenomenological theory would suggest, too, by its intentional orientation. This means that it is oriented toward something (which can be an object, event, theme, concept, etc.) that the sign designates as such-and-such. The sign thus “discloses” something in a specific way. This involves, in turn, a consciousness for which the sign takes on some meaning. Simil- larly, in the act of communication, it presupposes another consciousness for which the sign communicated becomes meaningful. In both cases, the symbolically mediated thought is (usually) directed toward that which is represented, not toward the representa- tional medium as such. There is thus no indication or implication of an inner theater of ideas, nor is consciousness oriented toward ideas. It is not. By thinking in signs, conscious- ness is oriented toward entities that are named or designated. At the same time, the “objects” are disclosed in some specific way, and it is this “sense” that creates the shared meaning among those thinking or communi- cating by means of signs. The concept of the sign involves thus a shared meaning, one that discloses something as something—that is, intentionally—in a specific way for both speaker and hearer.

Charles Sanders Peirce’s definition of the sign will now be used as a testing ground to further explicate our intuitive introduction of the sign. It will be immensely useful for showing that a philosophical conception of the sign can be made fruitful for our purpose. The fact that Peirce is able to articulate the sign as an entity that entails both an insurmountable relation to subjective consciousness and that is oriented toward something that it repre- sents is decisive here. Indeed, Peirce’s definition captures its most basic dimensions in one single formula: “A sign… is something which stands to somebody for something in some respect or capacity” (Peirce 1955: 99). This condensed formula can now be spelled out in its various aspects. In total, we can distinguish four important dimensions of the sign.

- First, the sign is *something*. The sign is thus designated in its material reality: it is something— it has itself a sensuous con- crete being. However, as we will later see, this materiality involves from the very beginning a mental or intentional structu- ration. Whether we take the phonetic sign or the written letter, the organization of
the material will only be the basis for meaningful expressions if it is integrated into a system of differences that assigns the mere material a significant spot.

- Then, the sign is oriented to somebody. As Peirce explains: “It addresses somebody; that is, it creates in the mind of that person an equivalent sign, or perhaps a more developed sign” (Peirce 1955: 99). What Peirce calls the interpretant of the first sign indicates three things at the same time. A thought can only be captured if it is mediated (or interpreted) by a sign. A sign is intrinsically related to someone else to whom it conveys some meaning; sign-use is communicative in its origin. Finally, the very idea of a sign involves, as being addressed to somebody, intentional consciousness. The sign must be understood as conveying meaning to someone, but as such it involves being a sign for a consciousness that understands its meaning. In other words, the intrinsic relation between symbolic meaning and intentional consciousness is here clearly marked. With this aspect, Peirce captures what we have called the reflexive dimension of symbolic mediation.

- Furthermore, the sign stands for something. This definition rules out—or at least does not involve—any reference to an inner mental sphere, to self-evidently present ideas or cognitions. The sign is seen as referring to the objects designated, be they real objects, other persons, inner feelings and mental states, or themselves symbolic expressions. It this “referential dimension” that many thinkers have thought essential to the definition of the sign as such, designating its nature as “something that stands for something else,” or as a “quid pro quo” entity.

- Finally, the sign stands for something to someone else in some respect or capacity. Here, Peirce indicates the crucial Fregean difference between sense and reference, the fact that the disclosure of an object as something involves a certain perspective, a conception or “sense” that is logically distinct from the existence of the actual object. The sign “stands for that object, not in all respects, but in reference to a sort of idea… “Idea” is here to be understood in a sort of Platonic sense, very familiar in everyday talk” (Peirce 1955: 99). Objects are not just directly given through signs, but are disclosed, represented so as to articulate and express a specific aspect or dimension of their being.

Peirce begins by accepting the reality of the sign as a symbolic marker. He then lays out (in points 2 and 3) his view that the sign will always entail the reference to both an intentional consciousness and to an object of reference. The really interesting and challenging issue, however, is introduced in point 4, where Peirce asserts that the symbolic relation will always represent something for someone in a determined perspective. This perspective, or cognitive perspective, may now be taken to introduce a mediating sphere between consciousness and reality such that the intentional reference to the thing or subject matter is made impossible. Indeed, it might seem that Peirce’s talk of “ideas” in a Platonic sense seems to recreate the very problem of the Cartesian philosophy of mind, i.e., that the intentional consciousness that is taken to refer to the world is indeed fully encapsulated into its own inner mental sphere of ideas.

One of the most challenging problems in the philosophy of language is how to reconcile the intentional objectivity of the sign with the perspectival dimension of disclosure that is involved in the understanding of the object as something. The acknowledgement of the intermediary understanding seems to reintroduce precisely the inner sphere of mental ideas that the linguistic turn was about to leave behind. What emerges as a challenge is how to show how the symbolic relation is capable of doing justice to both sides of the epistemic relation.

How can the meaning of the materially given sign (the sign is something!) be understood, if any direct grounding of meaning in either the reflexive consciousness or in the reference to some external reality must be excluded, because the sign contains a “sense”? This is a challenge because the understanding of any sign—and only thereby is understanding itself possible—means that the interpreting consciousness cannot draw directly or immediately on its own inner ideas or its own external experiences to interpret the sign (or thought) of another (or even, we might suggest, of itself). It must necessarily first be dealing with the sense, i.e., be concerned with the intermediary perspective or sense that is attached to the symbolic expression. In sum:

Thought as symbolic synthesis: Cassirer’s overcoming of Cartesianism

A systematic look at the contribution of Ernst Cassirer and his philosophy of symbolic forms will enable us to put this project into focus. Cassirer’s work perfectly situates him at this point in our inquiry, since his Kantian roots place him, on the one hand, in the tradition of the philosophy of consciousness, while his emphasis on the essential mediation of mentality through symbolic forms, on the other hand, make him the first modern thinker to place the symbolic mediation of mind at the core of a systematic cultural philosophy (Cassirer 1955a; Krois 1987; Friedman 2000).

Indeed, Cassirer represents an original fusion of Kantian and Hegelian motives, inasmuch as the mind is essentially dependent on symbolic forms for its capacity for self-realization. This project is realized, in a very novel turn, in the claim that the very structure and
essence of thought depends on the symbolic medium. Spirit is not conceived as pre-existent. Rather, the very root of mental life is already seen as requiring a medium of articulation, a symbolic form that allows for the specification, crystallization, and objectification of whatever the mental content might be. The Hegelian motive thus shows itself in the argument that spirit has to realize itself, that it has to be manifested — the manifestation, to be sure, now to be understood as achieved in the medium of symbolic forms. The Kantian legacy, however, remains alive insofar as the specific symbolic forms are seen as mediating categorial structures, as “frames of mind” (Gardner 1993) that structure all possible mental content along the pre-established lines of symbolic forms. In this regard, the original Kantian focus on cognition, even if broadened to include the transcendental grounds of moral and aesthetic judgment, proves too narrow. The new focus on symbolic forms makes it obvious that the “life of the mind” realizes itself in a pluralism of symbolic world-perspectives, including language, art, myth, religion, science, law, and the state. Perhaps Cassirer’s deepest philosophical claim consists in his critical reconstruction of the division between subject and object from within the symbolic mediation itself. According to Cartesianism, “subject” and “object” are seen as self-contained entities that stand opposed to each other, and that accordingly require some “theory of knowledge” of how one gets from within one entity or domain (the inner sphere of the mind) out toward the other entity or domain (the object in the “real” world that one claims to know or talk about). In opposition to this, Cassirer reverses the relation between the categories as mediating between the pre-existing mind and the pre-existing thing, by suggesting that the medium — the symbolic form as the enabling source of any categories of understanding — comes first. Instead, the symbolic form is the basis from which some subjective pole (or self-consciousness) and some objective pole (or thing) can alone emerge as later differentiations within the symbolic medium itself. Subject and object are thus not separated because these “poles” only emerge from within the symbolic mediation, and the respective symbolic forms create whatever we can know as the world. But in order to do so, we are thus already in a medium that connects us with and grounds us in the world. The symbolic relation is prior to the relation of a subject towards an object. Thus, no global skepticism is possible.

By thus overcoming the Cartesian split by means of reversing the order of explanation (or grounding), Cassirer in addition multiplies the worlds that we are able to experience and from within which subjects can know or relate to objects. If the symbolic mediation comes first, and subjective and objective differences come second, the reality of the symbolic form becomes the root of all experience and “truth.” Yet what must now be analyzed are the symbolic grounds on the basis of which the subject/object relation – or better: relations – is able to emerge. And now it can be shown that there is not just one symbolic form, but a plurality or multiplicity of “ways of world-making,” as the American philosopher Nelson Goodman (1978) calls his similar approach. Cassirer thus arrives at what we can call “symbolico-mental property-pluralism,” since he still maintains the unity of the mind as a source – an “original creative power,” as he puts it – while at the same time emphasizing that the actual realization of intentional thought implies a pluralism of symbolic world-disclosures.

The mind is taken to display its unity in a basic capacity for the symbolic function, while the actualization of that function involves the cultural and historical differentiation into diverse cultural forms (or fields). All forms have in common that they are symbolically mediated spheres of thought, but their specific intentional orientation differs markedly. Indeed, such differentiation is precisely what creates the diversity of cultural life, which is thus defined by the shared meaning of mental acts within each field and their basic difference with regard to their essential logical orientation in others. Intentional acts in art, for instance, are oriented toward beauty (or novelty, creativity, challenge), in religion toward faith, revelation, spirituality, in science toward truth and validity, in law toward justice, etc. The symbolic formation in each field allows for a precarious inner stability and identity of meaning of all acts uttered and displayed in each field. In our context, Cassirer’s approach is relevant because of the attempt to overcome Cartesianism on the very grounds of the philosophy of mind. In other words, the move to a semiotic version of the linguistic turn is not supported by (in themselves important) arguments concerning the communicability of ideas, the problems of reference or other minds, or the sharedness of conceptual meaning. Explaining these aspects of linguistic meaning plagued the Cartesian-empiricist philosophy from its inception in Locke (Taylor 1992). Rather, the identity of mental content itself is taken to depend on symbols, or rather, symbolic forms. As we will see, however, Cassirer’s strength proves finally to be his weakness: his philosophy of the symbolic mind remains too closely tied to a conception of the original intentional consciousness, and thus fails to capture essential aspects of a new viable theory of meaning. Yet the two arguments that he presents in favor of the symbolic mediation of mental content point represent a highly original contribution of an internal overcoming of Cartesianism based on an analysis of the possibility of synthetic and articulated thought. Cassirer presents two arguments that show that the intentional understanding of something as something — the essence of consciousness as being conscious of some-thing — can only be achieved on the grounds of symbolic mediation.

1. The argument from symbolic synthesis

Cassirer makes clear that it is language — or rather the sign as a name — that alone is capable of creating definite and permanent meanings. Critically oriented at an empiricist conception of the mind as a stream of sensory impressions (but also valid, by the way, for the phenomenological conception of the mind), Cassirer takes the symbol to fulfill the synthetic function necessary for identical (“ideal”) meaning. Ironically, it is a sensuous mode (a phoneme, a letter, a code) that allows the ideal meaning to stabilize, to crystallize, to fix its identity by “naming” something as something, and thus to bind the stream of differentiated and momentary states of awareness to an enduring and permanent idea. Permanent meaning, which alone is meaning, is thus possible because the mere psychological content acquires a “new permanence” by means of its conceptual stabilization in a symbolic form.

“In the immanent development of the mind, the acquisition of the sign really
constitutes a first and necessary step towards knowledge of the objective nature of the thing. For consciousness, the sign is, as it were, the first stage and the first demonstration of objectivity, because through it the constant flux of the contents of consciousness is for the first time halted, because in it something enduring is determined and emphasized. No mere content of consciousness as such recurs in strictly identical form once it has passed and been replaced by others. Once it has vanished, it is gone forever as what it was. But to this incessant flux of contents, consciousness now juxtaposes its own unity and the unity of its form... Through the sign that is associated with the content, the content itself acquires a new permanence. For the sign, in contrast to the actual flow of the particular contents of consciousness, has a definite ideal meaning, which endures as such” (Cassirer 1955a: 89).

The cognitive construction of meaning is here theorized in terms of the Kantian concept of synthesis, which in turn is to be understood in terms of the Kantian consciousness of the sensuous thought. Cassirer thus overcomes psychologism by means of a theory of the semiotic binding of sensuous thought.

2. The argument from symbolic articulation

Closely related is the fact that the fine-grained articulation that defines human consciousness cannot exist apart from the symbolic medium. The basic idea is that the differentiations of thought require a subtle structure to articulate conceptual and meaningful differences—a thought fully developed in later semiotics and structuralism. For Cassirer, the individuation of mental content is thus based on the articulation of that content as such-and-such content, which is not conceivable without language (see also Davidson 2001; Searle 1983, 2002). The identifying act of a name, made possible by the sign (as reconstructed just before) creates a category—say red or chair—which can now be applied to infinitely many new occurrences of the same feature or thing. But this in turn makes it possible to refer to red things in the past, in future contexts, in conditional or hypothetical settings, etc. There is a level of complexity in the expression of human thought that is unconceivable without language and its possibilities.

Accordingly, thought requires symbolic forms to articulate itself as higher-order thought. Language is thus a crucial factor in the “progressive process of determination” of meaning, and not just a cloak later added on to pre-established conceptual content in order to communicate.

“For the sign is no mere accidental cloak of the idea, but its necessary and essential organ. It serves not merely to communicate a complete and given thought-content, but is an instrument, by means of which this content develops and fully defines itself. The conceptual definition of a content goes hand in hand with its stabilization in some characteristic sign. Consequently, all truly strict and exact thought is sustained by the symbolics and semiotics on which it is based” (Cassirer 1955a: 86).

For example, the complex level of understanding that is achieved by physics would not be possible without the symbolic articulation of mathematical laws that truly define the object of physical science: “Without the universal signs provided by arithmetic and algebra, no special relation in physics, no special law of nature would be possible. It is, as it were, the fundamental principle of cognition that the universal can be perceived only in the particular, while the particular can be thought of only in reference to the universal” (Cassirer 1955a: 86). Even to identify a single object as an object—say a chair as a chair—requires both a concrete sign in which the fleeting moments of sensuous “chair-perceptions” are halted and synthesized to be those of a chair and, similarly and in the same act, the category of “chairness” that defines or subsumes those sense-data under the name “chair.” To achieve this is the function of the sensuous/ideal unity of the name-label “chair.” The linguistically grounded concept-formation thus allows for both the identification and articulation of thought, for the fixation of mental content and for the inner specification regarding the meaning of such content. But how exactly is the differentiation of subject and object—or better: the subjective and the objective poles—from within the grounds of symbolic mediation possible? How can we conceive of the emergence of the subjective and self-conscious “I” and the objective “it” (the Fichtean “non-1”) as moments or products of the symbolic function?

In order to reconstruct the emergence of reflexive consciousness from the process of symbolic mediation, I will present an ideal-typical argument based on Cassirer. It assumes, as does Cassirer if only in a heuristic fashion, that the mind is initially confronted with undefined and inarticulate sense impressions.

1. We assume thus a first level—a 0-level—at which consciousness receives undifferentiated impressions through the senses. Cassirer does not believe that this level is “real” as an experience, siding here with gestalt-psychological or phenomenological studies. But it serves nonetheless as a conceptual background for his thinking.

2. The fleeting moments of particular psychological experiences are then fixed and defined by acts of naming, as we reconstructed above.

3. The naming procedure makes the fleeting and singular acts “repeatable” in identifying them as the same experience (say of “red” or of seeing “a chair.”)

4. The sensuous-semiotic fixations of those fleeting moments in repeatable, reiterative names create concepts or ideal meanings. The vocal (or pictorial or written) gesture acquires an ideal meaning as it becomes an “entity” capable of expressing all possible similar events and objects. We can now speak of something as “red” or “a chair;” we can refer to such categories and things in past, present, future, hypothetical, and conditional states.

5. This move makes memory possible, because the subject can now refer to a content that it possesses and to which it can “return” in order to reflect upon it again.

6. This concept of memory thus transcends the idea of simple repetition and makes possible the attitude of a reflexive consciousness, since the content that is reproduced in the mind becomes available for that mind as a self-conscious object of thought.

7. We arrive, therefore, at a reconstruction of the difference, or better: differentiation between subject and object as emerging from the shared medium of symbolic thought. This is because it is this medium that makes possible the reflexive awareness.
of mental content in the context of memory, which in turn makes possible the awareness of "me" or an "I" for which this content, as an object of consciousness, exists.

The decisive move in the argument consists of the reference to memory as an essential step toward achieving a specific sense of self-consciousness:

“The concept of ‘memory’ thus takes on a richer and deeper meaning. In order to remember a content, consciousness must previously have possessed itself (sic) of that content in a way differing from mere sensation or perception. The mere repetition of the given at another time does not suffice; in this repetition a new kind of conception and formation is manifested. For every ‘reproduction’ of a content embodied a new level of ‘reflection’. By the mere fact that it no longer takes this content as something simply present, but confronts it in imagination as something past and yet not vanished, consciousness, by its changed relation to the content, gives both to itself and the content a changed ideal meaning” (Cassirer 1955a: 90).

It follows that the radical epistemico-ontological separation of the subjective inner sphere and the objective outer world are untenable:

“The limits of the ‘subjective’ and ‘objective’ worlds become for the first time really clear. One of the essential tasks performed by the critique of knowledge is to ascertain the laws governing this delimitation in the purely theoretical sphere, where it is effected by the methods of scientific thought. This [symbolic-constructivist, HHK] critique shows that the ‘subjective’ and ‘objective’ were not strictly separate spheres from the very beginning, fully defined in content, but that both became defined only in the process of cognition and in accordance with its methods and conditions... Seen in this context, the way in which we apply the conceptual opposition of ‘subjective’ and ‘objective’ in giving form to the world of experience, in constructing nature, appears to be not so much the solution to the problem of cognition, as its perfect expression” (Cassirer 1955a: 90–91).

While Kant and the Neo-Kantians see the distinction between the subject and object, and the respective creation of objective categories, as emerging from within the subject, for Cassirer this distinction is itself the effect of the symbolic process that entails both aspects as its internal differentiation. Yet in order to be fully convincing as an overcoming of the Cartesian subject-object split, Cassirer in turn has to show that the symbolic relation is one that bridges or mediates consciousness and the world. In other words, he has to make sure that the symbolic relation is prevented from collapsing back into an inner mental process of cognitive construction that in itself is possibly severed from real world groundings. It is a major step forward to ground self-consciousness in symbolic mediation because now we can relate it to a sphere of symbolic meaning that transcends the Cartesian encapsulation of ideas into an inner mental sphere. Yet we now also need a theory of symbolic meaning that presents us with the real world grounding to which Cassirer’s theory of the mind as a symbolic construction provides an essential building block. After Cassirer has convincingly shown that consciousness as symbolic mediation, we will now have to inquire whether his “Philosophy of Symbolic Forms” can present us with a convincing theory of meaning.

### Symbolic meaning and the limits of semiotic idealism

It remains to be seen whether this approach can fully account for the intricate dimensions of meaning constitution, which implies an explication of the semiotic structure of language, the relation between consciousness and language, the relation of linguistic units (statements or speech acts) to non-linguistic reality, and the social or intersubjective nature of language. In the course of the diversified forms of the linguistic turn, in both so-called “continental” and “analytic” thought, a different emphasis has been given to each of those, respectively. Arguably, a viable view of symbolic meaning has to give some account of each. How does Cassirer’s project fare in this regard?

We have seen how Cassirer succeeds in showing that symbolic mediation (via the functions of synthesis and articulation) is required for intentional consciousness. Yet symbolic mediation cannot itself be mental or (merely) cognitive. We need a basis prior to the intentional understanding of language from which such understanding (which entails and internally differentiates subject/object) is to emerge. Cassirer’s philosophy presents us with two such sources, one individual, the other collective. On the individual plane, Cassirer believes that we find in emotional expressions a paradigm for self-interpreting entities that are what they mean and mean what they are. We find here an “experience” of symbolic unity, expressed by the “immediate” understanding that agents possess when confronted with basic emotional expressions.13 On the collective plane, such an immediacy of meaning is provided by the symbolic form of myth (Cassirer 1955b). Both have in common that prior to the subject-object separation, there exists a unified experience of symbolic meaning where the word and the thing are one. Something appears as meaningful, but is not reflexively interpreted – it does not exist, as phenomenologists and Heidegger (1962) may say, as an apophantic As. Rather, it exists as a hermeneutic or symbolic As, i.e., it gives what it is in the act of understanding by articulating the structure of the object itself, prior to being cognized as an object.14

We have seen that Cassirer argues that both modes are the most basic exemplifications of the symbolic function. This certainly helps us understand that it can bridge the gap between mind and world. And yet, it does not as such already entail a solution to our problem, but rather expresses the need for a theory that can explain which differentiations between subject and object-positions are possible. To see this, we refer to a deep and ultimately unresolved ambiguity in Cassirer’s conception of the symbolic function. Indeed, at the core of Cassirer’s project lies a peculiar yet highly instructive paradox. On the one hand, Cassirer emphasizes the intrinsic dependency of thought (and mental life in general) on the medium of symbolic expression. Indeed, it is this move that allows him to overcome the narrow orientation toward a merely cognitive-scientific comprehension of intentional meaning, allowing for the plurality of symbolic forms or worlds. By reconceptualizing thought as a multi-various domain of symbolic expressions, he achieves
a unique connection between the symbolic function and the diversified fields of cultural self-expression. On the other hand, however, the plurality of fields is understood as ultimately emerging from a basic source of creative human imagination, from the ever-active origin of the mind. The "manifold manifestations of the mind" are taken to emerge from an "original imaginative power" (Cassirer 1954: 88), an "original formative activity" (Cassirer 1954: 90) that underlies and grounds all specific symbolic formations. The paradox shows itself in a methodical circularity, which points to the unavoidable hermeneutic dimension of all human cognition: The symbolic forms are taken to be ultimately grounded in the original power of the mind, while the mind itself – if we take the arguments concerning the articulation and fixation of meaning seriously – can only be articulated through the pluralized media of its objectified forms. It is thus symbolic forms, in their diversity, that exemplify the reality of cognition, which nonetheless are to be attributed to an original, foundational, and underlying origin in the mind.

What we now need to do is to inquire whether this foundational conception of a mental source of meaning limits a full understanding of the intertwinedness of language and thought in Cassirer. Indeed, by exposing the limits of Cassirer’s symbolic theory, we derive a systematic outline of the following reconstruction of the philosophy of language. We will follow as our guide the four basic issues that we outlined at the outset of this section – the account of symbolic medium, consciousness, reference, and intersubjectivity – and analyze each with regard to its treatment in Cassirer. At stake is whether the primary role that Cassirer still attributes to an original power of mental activity conflicts with an uncompromised account of these domains of symbolic meaning.

1. Intentionality and symbolic power

A first issue arises with regard to the alleged autonomy of meaning–constitution that Cassirer asserts with regard to meaning and that may ultimately be attributed as emerging from an original formative power of the mind. While Cassirer emphasizes that the mind needs to symbolically express itself – it literally only exists in the very medium of its symbolic forms – the idea of an underlying mind invokes a very strong conception of independence and freedom of consciousness. As we saw, meaning, in order to exist at all, requires its sensuous–symbolic expression. The sign fixes and articulates the content, and, as such, makes its ideal identity as mental (and not just psychological, empirical, private) content possible. Meaning must be "reproducible," that is the sign’s iterability allows mental content to be objective and shared. However, for Cassirer this move does not do away with the connotation of an original meaning-imbuing act, an act, furthermore, that he conceives to be absolutely free and independent from any external constraint.

"…the more clearly the particular cultural forms disclose their specific energy, the more evident it becomes that all apparent “reproduction” presupposes an original and autonomous act of consciousness. The reproducibility of the content is itself bound up with the production of a sign for it, and in producing this sign the consciousness operates freely and independently" (Cassirer 1955a: 90, my emphasis).

This formulation condenses, in a nutshell, the whole problematic of Cassirer’s fusion of language and thought. For it is not entirely clear whether the “original and autonomous act of consciousness” refers to a transcendental consciousness that grounds the possibility of the symbolic form as such, or whether it refers to each interpretive act of understanding that is undertaken by every individual speaker and hearer by using language. If, on the one hand, the “act of consciousness” is meant to designate a transcendental power prior to the linguistic or symbolic forms, we are back in the classic Cartesian conception of pre-linguistic mind that can spin its forms out of its inner cognitive cave. This is a conception of consciousness that Cassirer himself seemed to reject with his strong thesis of the symbolic mediation of mind. If, on the other hand, the act of conscious understanding refers to the individual speaker, the idea of the autonomy and independence of consciousness indicates a problem rather than a solution. This is so because if the interpretive acts of speakers and hearers exist in a sphere of meaning that is disclosed through shared symbolic forms, the individual act of understanding is not simply a free and autonomous endeavor. In fact, understanding and expressing meaning then presupposes a conformity and abiding by the general structures and rules that define the shared symbolic meaning in the first place. The relation between the speaker’s intentional meaning and the conventional, structural, or communicative meaning needs to be clarified.

Accordingly, the very idea of a symbolic mediation of mind involves a rethinking; in fact, a radical modification of what Cartesianism and idealism assumed to be a free and autonomous act of consciousness. If consciousness as such (the transcendental dimension) requires language, and if symbolic mediation thus prescribes a shared realm of meaning prior to each individual act of understanding, consciousness (understood as the concrete comprehensive act of an individual mind) is not free and autonomous. Consciousness rather depends on the shared symbolic forms that have emerged in the social and cultural practices in which situated speakers always already find themselves. In this regard, the symbolic forms set limits for individual speaker-minds: they redefine, structure, orient, and shape what can possibly become a mental content. If symbolic mediation is thus bound up with social and cultural practices, and if it predetermines the realm of possible thought for individual speakers, it will prove to be a dimension of power over the mind. The implication of power as related to the symbolic mediation of mind must thus be addressed (Foucault 1990; 1972; Hacking 2000).

2. Language and consciousness

The second issue arises with regard to the relation between natural language and intentional consciousness. As we saw, Cassirer takes the mind to depend on its symbolic externalization. The symbolic forms that are thus created include language, myth, religion, science, philosophy, etc. However, natural language is thus not granted a basic status, but appears to be merely one among other possible forms of human expression. This limitation of the role of natural language for the formation of symbolic forms in general seems problematic at least in three regards. First, the attribution of higher-order beliefs in general seems to depend on linguis-
tic structures. As Cassirer himself shows with regard to basic concept-formation, the naming-act crystallizes and synthesizes the sensuous flow of impressions. Concept-formation, and the fixation of meaning, however, is a basic requirement for all other symbolic forms, including myth, religion, science, etc. The very articulation of complex beliefs, which takes evidently place in those domains, makes use of the capacity to express those thoughts in language. The fact that we are unwilling to ascribe complex beliefs to infants at a pre-linguistic stage or to animals displays this basic intuition with regard to an in-principle function of language for higher-order human thought. Second, artificial languages such as sign-systems used in mathematics (which arguably create mathematic structures) are introduced and learned in a context in which agents already understand a language, in which they can coordinate their interactions so as to define the signs to mean what they want for them. This meaning, then, is fixated and articulated by the refined and specifically constructed sign-systems. But, as problems of interpretation or ruptures in the shared consensus regarding a scientific paradigm may show, the background context of a shared natural language is essential for the sustained existence of thus-constituted meaning.16 Third, so-called non-verbal languages, perhaps better called non-linguistic sign-systems, derive the full depth of their meaning to a large extent from their embeddedness in linguistically mediated contexts. Take, for example, the interpretation of the “language of painting,” in which figures are displayed in the medium of form and color. As the iconicographic method in art history made clear, interpretive access to the meaning of particular scenes (say the Jesus birth scene) requires an in-depth study of the textual background disclosing this scene at the time (say 13th century Italian Bible interpretation, when a Giotto or Cimabue painting is in question.17

The foregoing reflections are not meant to suggest that there are no a-linguistic systems of meaning, nor that all meaning can be reduced or fully translated into linguistic forms and expressions. Yet they suggest that there is good evidence that human understanding is essentially structured by the linguistic dimension, which in turn demands that we analyze the role of natural language for the construction of human thought in its full and pervasive function: “Being that can be understood is language” (Gadamer 1989).

3. Symbolic mediation and reference

The attractiveness of the linguistic turn is grounded in overcoming the cognitive encapsulation of the mind. Grounding the reconstruction of basic philosophical concepts in the symbolic relation (understood in a realist way) avoids Cartesian skepticism and solipsism. Yet the way in which Cassirer conceives of symbolic mediation – a way ultimately due to his roots in the philosophy of mind – counts precisely this advantage. Because Cassirer conceives of symbolic mediation solely with regard to mental content, while conceptualizing this content in terms of ideas (only now fixed and articulated through signs), a new symbolic idealism emerges from his position.

One of the intrinsic dimensions of the sign is its “intentionality.” We argued that this inheres in the sign. The “directedness of a mind” is not toward its own inner mental representations. Rather, when we communicate with someone about something, we are usually oriented toward something in the world. Symbolic intentionality thus points intrinsically beyond the boundaries of language itself. Certainly, it mediates our understanding by shaping and articulating our experience, but it nonetheless remains an experience of something that is beyond the mediated realm of symbolic expressions. Cassirer, however, loses this sense of an intrinsic openness of language toward the world by conceptualizing symbolic mediation as the creation of holistic and self-contained worlds – symbolico-cultural constructions that create their own inner network of references and as such attain a unique form and autonomy of meaning. The symbolic mediation is not seen as a mind-world-interactive construction of meaning, but rather as the emergent constitution of self-referential symbolic worlds.

Cassirer supports this vision by adopting Heinrich Hertz’s conception of the world of physics as an entirely self-contained universe of meanings. True, the physicist seems to relate to sense-perceptions that he has to organize in order to create a law-like interpretation of nature. But this does not mean that “physics” remains in any significant sense tied to such entities or events.

“The concepts with which [the physicist, HHK] operates, the concepts of space and time, of mass and force, of material point and energy, of the atom or the ether, are free ‘fictions.’ Cognition devises them in order to dominate the world of sensory experience and survey it as a world ordered by law, but nothing in the sensory data themselves immediately corresponds to them, yet although there is no such correspondence – and perhaps precisely because there is none – the conceptual world of physics is entirely self-contained” (Cassirer 1955: 85).

To be sure, Cassirer is fully aware that the intentional direction of meaning points beyond its own self-set boundaries: “…all these symbols lay claim to objective value. They go beyond the mere phenomena of the individual consciousness, claiming to confront them with something that is universally valid” (Cassirer 1955: 88). However, since according to Cassirer this claim is itself the effect of a symbolic function, it will ultimately be fleshed out only through a symbolically self-contained form of meaning, one that is accessible to all minds in principle. In this regard, the natural sciences do not achieve their truth through some “correspondence” to objective states of affairs: “the conceptual world of physics is entirely self-contained. Each particular concept, each special function and sign is like the articulated word of a language meaningful in itself and ordered according to fixed rules” (Cassirer 1955a: 88). Galileo’s metaphor of the “book of nature” is to be taken literally; the symbolically “self-enclosed” worlds achieve their objective meaning, it seems, only by creating such self-referential perspectives of meaning, even or especially in the case of the natural sciences. Ultimately, scientific theories are universally valid, but only as universally accepted fictions.

Again, Cassirer’s project proves to point to a problem rather than provide a solution. While he correctly urges us to transcend a naïve correspondence theory of knowledge, the claim of the symbolic mediation of mental content cannot lead straight back into a Neo-Kantian idealism of symbolic forms. The intentional openness of the sign toward the world needs to be reconstructed according to its phenomenological meaning, that is, according to its intentional-referential orientation toward extra-linguistic entities.
4. The intersubjectivity of language

Our initial reconstruction of the meaning of a sign suggests that its social use and origin are essential components of its structure. The idea of shared meaning associated with sign-use is not just an arbitrary addition, adding on to language, but entailed in its very identity. Yet Cassirer’s focus on the symbolic mediation of mind misses what this involves. Symbolic forms spin themselves out of the “original formative power” of the mind, only to create new trans-subjective modes of meaning into which individual speech acts and their communicative interaction – exactly how we are never told – are taken to somehow fit. How individual speech acts are woven into an intersubjective, and thus normative, context of social relations remains entirely untheorized.

Indeed, Cassirer conceives of symbolic mediation fully in terms of the Subject-Object model. This becomes clear in his take on Humboldt, whom he credits with having first understood the essential cognitive function of language for consciousness. Quoting Humboldt, we see that “just as the particular sound mediates between the object and the man, so the whole language mediates between him and the nature that works upon him from within and without” (Cassirer 1955a: 92).

According to the symbolic structuration thesis developed earlier, language is here seen as a formative power regarding inner and outer nature. Cassirer comments on this section by emphasizing the reflexive gain that such mediation entails: “In each one of its freely projected signs, the human spirit apprehends itself and its own formative law. And this peculiar interpenetration prepares the way for a deeper determination both of subject and object” (Cassirer 1955a: 92). The point with regard to the inherent reflexivity of language is well taken, as we saw above. Indeed, it is this linguistic reflexivity that defines the dimension of language and its use became central from the start. If we read the linguistic constitution of mind, nonetheless oriented at a paradigm that conceives of consciousness as an Ur-source of experience and meaning constitution, the shaping power of language is not yet fully acknowledged or theorized. First, Cassirer sticks to the rhetoric of autonomous acts of meaning creation, while the linguistic-symbolic turn would have suggested a displacement of the power of the subject over language. Second, consciousness is seen as the source of a plurality of equally basic symbolic forms, instead of recognizing the fundamental role of linguistic understanding for all possible human experience. Third, Cassirer pushes the idea of a symbolic construction of reality toward an idealistic conception of reality, which loses contact with non-linguistic reality and the language-world relation. Fourth, Cassirer entirely skips an analysis of the intersubjective dimension of language use, which means that the relation between normativity and social rules is missing from his theory.

These issues have been at the center of various forms of the linguistic turn. The relation between language and thought with regard to the issue of symbolic power has been at the core of the semiotic approach of de Saussure and his successors in structuralist and post-structuralist linguistic and social theory. Here the speaker’s interpretive agency and symbolic structures are mostly considered an expression of social power and, as such, constrain the individual consciousness. Similarly, the analytic approach is mostly concerned with truth and reference, taking its cue from Gottlob Frege’s ground-breaking distinction between sense and reference, but eventually leading towards a pragmatic account of meaning in Quine and Davidson and their followers. Finally, in speech act theory, evolving from the philosophy of late Wittgenstein, the normative and practical dimension of language and its use became central from the start. If we read the linguistic turn in this way, it entails the following three assumptions:

(a) The abandonment of any reference to a foundational or constituting consciousness, since shared social practices of language use become the starting point of the philosophical analysis of meaning.

(b) The abandonment of any attempt at a naive, direct, or metaphysical realism, since all experience is symbolically mediated and thus relative to the language games in play.

(c) The acceptance of the intentional nature of consciousness, which involves the referential relation to some object or subject matter of experience.

Cassirer has given us proof that the symbolic reconstruction of consciousness is both necessary and possible. Even though his continuous reference to the mind as an original source of meaning prevents him from pro-
providing us with a consistent theory of symbolic meaning, Cassirer has shown that such a theory will have to entail a conception of intentional consciousness as symbolically mediated.

Notes

1. In an important study aimed at reconstructing the analytic/continental split in contemporary philosophy, Michael Friedman notes: “It must certainly be admitted, therefore, that his synthetic and conciliatory approach to both philosophical and political questions makes Cassirer a much less striking and dramatic figure than either Carnap or Heidegger. Those interested in finally beginning a reconciliation of the analytic and continental traditions, however, can find no better starting point than the rich treasure of ideas, ambitions, and analyses stored in his astonishingly comprehensive body of philosophical work” (Friedman 2000: 159). I owe this quote to Jared Warren.

2. It has nonetheless been associated with the semiotic or structuralist paradigm. This interpretation of a structural semiotics as objectivistic found its source in the Anti-Cartesian stance of this approach. It did indeed lead to a wide-spread methodological self-misunderstanding of semiotically influenced thinkers. However, we will show that the reconstruction of meaning as founded upon an implicit holistic background structure never loses touch with the intentional understanding of speakers as agents.

3. This assumption is so basic that it often goes unnoticed, even in semiotic analysis itself. The very reflection on this fact is equally made possible by language and its inherent reflexivity, which, again, is often not emphasized with sufficient rigor.

4. The integration of a singular sign in a symbolic context – whether it is sentence structure or a holistic system of a wider context – anticipates later developments, esp. the semiotics of de Saussure.

5. Indeed, according to the full-fledged theory of the symbolic construction of thought, we do not just need to interpret the signs as used by someone else, but actually need to constantly interpret our own use of signs. Thinking is the self-interpretation of our own use of symbolically mediated ideas, or an internal interpretive dialogue between signs and our understanding of them.

6. This broadening of the focus of Immanuel Kant’s threefold transcendental critique – including natural-scientific, moral, and aesthetic judgments – has generally been a motive of neo-Kantianism. Especially, the so-called Southwestern School (for instance, Heinrich Rickert) attempted a quasi-Kantian grounding of the human and social sciences, which lack any reflection in Kant himself. Cassirer continues and enforces this line of thought, but also lifts it to a new plane through his theory of the necessary symbolic mediation of thought.

7. The symbolic turn in neo-Kantianism can take two forms: either of pluralizing the validity claims along the lines of Kant’s three critiques; or of pluralizing the domains – the symbolic worlds – within which validity claims can be made. Cassirer himself shifts between two positions, of which the first is not identical in content to the Kantian threefold division. In his major work, “Philosophy of Symbolic Forms,” Cassirer suggests a systematic threefold structure that sees emotional, representational, and self-referential symbolic expressions as the basis for the respective cultural symbolic forms of myth, language, and science. In other contexts he defines art, the state, and law as additional symbolic forms, thus indicating a more open-ended and empirical notion of the plurality of existing symbolic forms. In both cases, however, Cassirer has made the essential move to ground the intentional consciousness in a prior situatedness in the symbolic function, and to understand the objective content of subjective experience as essentially mediated by diverse symbolic forms.


9. We will later problematize Cassirer’s adherence to a unified mind as the basis of all multiple worlds, while similarly maintaining that consciousness is only possible as a reflexive effect of the symbolic function.

10. The sign is always to be understood as embedded in a sentential context, because only as a statement about something can a name be considered a name. “Chair” is nothing but a sound-thing, unless it is used in a (however implicit) speech act, “(this is) chair.”
11. The concept of the sound as a sensuous medium of thought that allows for the stabilization of the fleeting perception is strongly influenced by Wilhelm von Humboldt (1988), whom Cassirer discusses with high approval.

12. Cassirer’s argument may seem like a semi-otic version of Sellars’s famous argument against the “myth of the given.”

13. For a critique of the immediate understanding of emotions as basic and unmediated, see Köglér/Stueber (2000).


15. For a similar criticism of Cassirer, but without the intuitive reference to animal cognition, see H.-G. Gadamer, who charges Cassirer with missing “the universal function of language” that hermeneutics encounters in the pervasive and encompassing nature of human understanding. (Gadamer 1989: 403–404).

16. Such a reference to so-called “paradigm-shits” in the sciences do not compel us to accept some of the implausible interpretations of Thomas Kuhn’s thesis of incompatibility or incommensurability of paradigms. It is just when the scientific community has to decide on a theory change that great consequences that a more encompassing cultural pre-understanding can be shown to play a certain role. Recently, the SSK (Studies of Scientific Knowledge) approach has expanded on the role of such external factors, which is not to say that they alone or fully decide theory choice. See Pickering (1995).

17. The art-historical approach of iconography was developed by Erwin Panowski, and has even influenced the social theory of Pierre Bourdieu.

18. The Humboldt quote is from the famous “Einleitung zum Kawi-Werk,” see Humboldt (1988).

19. The complete absence of this dimension of language or meaning also becomes apparent in the context of the genetic question of the origin of language. Cassirer seems to be aware of only two possibilities, i.e., that the first linguistic human signs derive: either (a) from an emotional expression of inner states, as in “ouch” or laughter, which is then turned into signs; or (b) from the imitative articulation of externally experienced sounds, as in “kikeriki” or “wind” (See Cassirer 1955a: 92ff.). But by seeing only the emotional or the onomatopoetic sources of language origin leaves out the third option, namely that the first signs emerged as a tool to enable social cooperation, which might have been essential for the survival of a physically weak animal, the human. The genetic question of language origin is not of major interest, however, and was mentioned here only to illustrate the absence of the intersubjective dimension from Cassirer’s thinking.

References


Received: 1 May 2009
Accepted: 7 July 2009
Are Enactivists Radical?

Dawid Lubiszewski
Nicolaus Copernicus University, Poland

Daniel Hutto is professor of philosophical psychology at the University of Hertfordshire. His interests focus on understanding human nature. *Radical Enactivism: Intentionality, Phenomenology and Narrative* illustrates Hutto’s work and interests. In this book, he gives new insights into intentionality, emotions, and enactivism. Hutto’s approach challenges many different perspectives and Hutto is not afraid of stating his ideas in a provocative way. He invited his philosophical opponents to contribute chapters to this book. The book consists of a central paper, *Unprincipled engagements: Emotional experience, expression and response*, written by Hutto, and six objections raised by famous cognitive scientists, complemented by Hutto’s answers to each of them. This book is recommended for anyone who is interested in cognitive science, especially for those who would like to get to know better the enactive approach or problems of philosophy of mind. The language of the book is not too technical and targets both novices and experts alike.

Is Daniel Hutto’s book *Radical Enactivism* radical? To answer this question it is necessary to address a different question: What is enactivism? In his introduction to the book, Richard Menary gives a definition that is a view opposed to more traditional theories. Undoubtedly it is correct to say that enactivism challenges the traditional ways of thinking about cognition and the mind. Menary also classifies different varieties of enactivism – from “traditional” (Varela, Thomson & Rosch) to “radical” (Noë, Hutto). All of them agree that the agent is “cognitive” if it has a body that is embedded in an environment with which it interacts. The “traditional” enactive approach uses the term “autopoiesis” to explain the interaction between the organism and its environment. However the “new, radical” version of the enactive approach does not need and does not use the autopoietic definition of an organism. Indeed, rejecting autopoiesis makes the “new” version of the enactive approach different from its traditional roots; however, I do not think that it makes it radical. It is not surprising that Hutto’s approach may appear radical if it is compared to the traditional and old-fashioned ways of thinking, but if we consider some details of Hutto’s approach, then it looks original, challenging, and fresh – rather than radical.

One of the main differences between the so-called traditional approaches and the enactive one is the rejection of the notion of “inner representations” by the enactive one. However, there are some types of representations that are accepted and used by enactivists. Menary draws on Wilson’s classification of representations, which distinguishes reactive, enactive and symbolic representations. Neither reactive representations – that is, reactions to the states of the world – nor enactive representations – reactions that involve bodily skills but do not involve conscious or reflective thought – carry any “content.” Both those types are accepted by enactivists such as Noë. Enactivists, however, criticize symbolic representations that are independent from the here and now. Moreover, Menary suggests that if we want to know what it is to be a representation, we can use the Peircean principle. This principle involves three components: (1) a vehicle has some inner or relational features that are salient; (2) the vehicle’s function of representing x feature is established by exploiting the vehicle by a consumer using its salient features; (3) when the representational function is recruited for a certain purpose, a genuine representation is established. The Peircean principle is very useful because it also explains how representation works in both natural and social environments, without saying whether representations are internal, external or distributed across the body and the world. Enactivists also avoid the “encoding” view of representations, which is shared by cognitivists. Still, Menary warns enactivists that the radical nature of the enactive approach may be lost, and that enactivism would then become nothing more than just another form of cognitivism. This would happen if they construed knowledge of the laws of sensorimotor contingencies as...
knowledge of the laws of symbolic representations of the encoding of those laws.

So what is special about Hutto's view? As Menary points out, Hutto’s approach distinguishes intentionality that is not understood as a property of a mental state or mental representation (i.e., the form of perceptual and emotional responses) from intentionality-involving truth conditions (i.e., the level of linguistically mediated thought). The first type of intentionality refers to biosemiotics (i.e., biosemantics without the semantics). Therefore, this type of intentionality is contentless and does not involve truth conditions and reference. Clearly, Hutto rejects the idea that contents and experiences are kinds of objects, for example mental objects. Feeling the experience of doing something does not involve any theoretical or conceptual objects, as it is guaranteed by directed enaction or imagined re-enaction. The view that he rejects is called “Object Based Schema” (OBS).

Basically, Hutto rejects such views, where experiences are thought of as the objects of selective attention. They should be thought of as actions done by or done to the experiencer. He says that instead of referring to a specific action as an object, it would be better to simply give descriptions of its character, such as “my feeling of pain in my arm.” For Hutto, experiences are nothing more than abilities and can be understood by our active engagement with environment. Even if, in general, his view may seem similar to the views of Alva Noë and Kevin O’Regan, it is different when it comes to details. He shares the idea that experiences are not objects and accepts the basic idea of O’Regan and Noë’s sensorimotor contingency (SMC) approach. However he also rejects some specific points in this account. His objections consider the notion of practical knowledge, which he finds problematic. Hutto’s approach avoids the “hard” problem of consciousness. The epistemological problem of other minds disappears in Hutto’s account. The traditional thinking in terms of the inner states of other people is replaced with a new theory – the theory of reacting and feelings as responses to specific expressions of emotion in others.

In the question-and-answer part, Hutto deals with objections made by seven cognitive scientists. It can be said that almost everyone will find something interesting and something that they can accept. However there are some concerns about Hutto’s account. In Feelings and objects, Erik Myin and Lars De Nul question Hutto’s objections to object-body schema and to the “hard problem” of consciousness, as well as his interpretation of SMC. They are not the only ones who find the rejection of OBS problematic. Anthony Rudd in Unnatural feelings, Tim Crane in Intentionality and emotion and Peter Goldie in Emotional experience and understanding ask Hutto for a clarification of his understanding of OBS. Rudd’s objections also touch upon Hutto’s definition of emotion. Peter Hobson’s point of view is similar to Hutto’s, but in From feeling to thinking (through others) he asks his colleague for further explanation about the problem of the other’s mind. Finally, Shaun Gallagher in The narrative alternative to theory of mind opens up discussion about the mentalistic theory of mind and narrative theory. Hutto’s reply to Gallagher introduces the “Narrative Practice Hypothesis” (NPH). NPH suggests that the ability to mindread is not biologically inherited from our hominid ancestors. Some of Hutto’s responses are longer than the article that they defend. On the one hand, his replies show that he understands the objections and he genuinely wishes to give a proper answer to each of them. On the other hand, those long answers may suggest that he was not clear and precise enough in the main paper. It is partially true because in a few of his replies he clarifies his ideas. Moreover, Hutto’s explanations prove that he knows the subject very well and has good arguments for his theories.

To sum up, Hutto’s approach is yet another one that challenges cognitivists’ ideas. What makes Hutto’s account special is his commitment to the rejection of content, a point where he becomes a real radical. He has an outstanding ability to look through any theory and find something useful. In other words, it is not just another book about enactivism but it is an enactive book for everyone written by an enactivist.
Breaking the Spell of Volition

Przemysław Nowakowski • Nicolaus Copernicus University, Poland

Our mental life is shaped by our acting in the world, in the process of coping with spatial and social problems. The mental is built in an interaction between the agent and the environment (Thompson & Varela 2001). But how is this process realised? What can we say about it? What is going on when it does not proceed appropriately? Is the whole sphere of the mental connected with action in the same way and to the same degree?

It seems that in the wide spectrum of what is traditionally conceptualized as the mental we can distinguish many components more and less intimately connected experientially to our actions. The feeling that we are the agents and controllers of our actions seems to be the component most directly connected with our actions. According to one widespread view, there holds a simple correspondence between action and the experience of action, thanks to direct perception. Contemporary research, however, challenges this popular idea (Haggard 2009). In Disorders of Volition we can find a rich and interesting body of research that shows how the subtle construction of our sense of will and agency can be broken. Some indication of the complexity of the issue of volitional action and its experience is provided in the introduction. Prinz, Dennett and Sebanz give us examples of several distinctions pertaining to volition:

1. How we realize our action vs. how the goal of our action emerges
2. The dynamics and the mechanics of action
3. Choice of actions and goals, and achieving them
4. Planning vs acting
5. Agency and authorship of action
6. Mental causation.

Their introductory character notwithstanding, these remarks already manage to convey some of the intricacy of the problem.

Moreover, if you accept the principle that any good theory of will should explain the pathologies of will. This principle points out the importance of testing any disturbance to determine – whether directly or indirectly – what the subject of the disturbance is (Can-guilhem 1991). This book is a very good introduction to research on the pathology of volition and a contribution to a better conceptualization of what it is that we call “will.”

Disorders of volition is comprised of five parts. Each part discusses a separate area of research:

1. Theoretical reflections on phenomenology, in particular on the sense of agency and intentions in action (Metzinger, Bayne & Levy, Haggard, Proust, Ainslie, Cohen & Gollwitzer)
2. Neuroscientific research on will, and the disorders of volition in schizophrenia (Jeannerod, Liddle, Spence & Parry, Frith)
3. The analysis of the relationship between volition and depression (Nitschke & Macewicz, Schneider, Jouvent & Fossati & Dubal)
4. Descriptions of the failure of will in patients with prefrontal lobe damage (Owen, Grafman & Krueger, Burgess & Gilbert & Okuda & Simons)
5. Finally, the relationship between will and substance abuse disorder (Bechara, Sayette, Hull & Slone).

Despite the suggestion at the beginning of this review that the chapters of Disorders of Volition are relevant to the philosophical question of the construction of free will, it must be said that this book relies heavily on specific experimental material, and no attempts are made to formulate any general conception of volition. Could this research be applied to constructivism? It seems so, but only in a more subtle way. Some of the chapters illustrate the process of the construction of that which we experience as our “will.” This kind of experience is a complex construct that is dissociable from action. Let us consider three examples.
Haggard looks into conscious intentions. His chapter (and 2009) is part of his long-standing research on volition. He points to three main features of conscious intention:

1. A sense of agency
2. An implicit, but important, difference between real time and perceived time in volitional and non-volitional action
3. A reconstructual and predictive account of intention and agency.

More specifically, a sense of agency is key to experiencing intention. But the sense of agency—which is the sense of being the agent of the originator and controller of action—is connected to temporal binding. This, in turn, is an effect present in volitional action, whereby for the same interval, the time between the cause and the effect is perceived as shorter in volitional than in non-volitional action; this mechanism binds causes and effects in our experience. In other words, this makes our experience of intentional action coherent.

The second element is an anticipational account of agency. That we perceive ourselves as agents is partly the effect of the anticipation of the consequences of our actions. This account has lately received experimental support (Haggard 2009). The sense of agency and, by the same token, of conscious intention, arises from the creation of models for the anticipation mentioned above. Another perspective is the so-called inferential conception (proposed by, among others, Daniel Wegner), which posits that the sense of agency arises from the interpretation of already instigated action. The two outlooks, however, are partly complementary rather than exclusive, and Haggard’s results are a weighty contribution to what is currently known about will.

Another important text is the one by Chris Frith. This seasoned researcher of schizophrenia reviews the classic studies on schizophrenia as well as his own ideas regarding this condition. However, he also adds a crucial element. Discussing two aspects of schizophrenia—(1) passivity experiences and (2) lack of will—he suggests that their explanation requires a reference to social factors. More specifically, the first element is present when the cognitive system does not distinguish between the actions of itself and those of another agent. The second is the effect of the inability to generate action suitable to the current social context.

This means that in addition to complex sensorimotor mechanisms, the construction of the will requires social coherence and a correct model of oneself as distinct from others. These elements are often neglected in the constructivist study of will. Unfortunately, it is not developed in detail in Frith’s text, either.

Still another important paper is that by Schneider, dealing with the problem of a disorder of volition in depression and how it comes about. Schneider interprets this disorder in terms of a model of action in which our actions are guided by our goals and the goals are assigned affective values; what is more, the effects of the actions are remembered. How is depression explained in this account? It is taken to be a disorder in which the subject does not initiate any action because they see the probability of achieving their goals as too low. Central to this disorder is the evaluation of the action as unfeasible, which is precisely what Schneider’s model accounts for. The lack of implementation of many actions—in the social dimension, also—is a frequent cause of negative affective consequences. This, in turn, has an impact on subsequent action: it generates only transient mood changes for small failures; but in the case of constant failure, it leads to depression.

The three examples above demonstrate how numerous and varied the factors are that influence the creation of the relatively frail construction of our will, as well as those that are responsible for its disorders.

Unfortunately, Disorders of Volition does not cover all the interesting aspects of volition. The problem of the indication of the role of social interaction in the neurocognitive research on volition is not developed in this book (Frith only mentions this issue in pages 243–244). But this perspective is currently present in literature (Sebanz 2007) and makes our understanding of volition and its disorders almost complete. This shortcoming, however, appears to be the only weakness of this book.

The volume, edited by Prinz and Sebanz, is interesting and worth reading as an introduction to the philosophical and cognitive neuroscience research on volition. It does “break the spell” of volition being the elusive question of will unavailable to scientific treatment, and indicates how ubiquitous this problem is in cognitive-neuroscientific research.

References


Recent books and articles related to constructivist approaches

For more information, visit http://www.constructivistfoundations.info/bib/

Purpose: Starting with the March issue of Constructivist Foundations, this section lists publications related to constructivist approaches – constructivism, second-order cybernetics, enactivism, non-dualism, biology of cognition, etc. – that have been published elsewhere since 2008, and which the reader of the journal might find interesting.

Content: The entries are ordered alphabetically and clustered according to their respective primary disciplinary backgrounds. The list of publications can also be found in the journal’s database at http://www.constructivistfoundations.info/bib/. The section will be regularly updated so that over time this will become a standard reference for constructivist literature.

How to contribute: To have your constructivism-related publications listed in this section, send an email to ariegler@vub.ac.be. Please format your list in the same way as the entries in this section. Add information about the disciplinary background and, optionally, a short abstract of one or two sentences (which should describe the link to constructivist approaches), and a URL that points to the full-text version. If you would like to review one of the books listed, write to ariegler@vub.ac.be.

Communication Science


“The author compares the development of constructivist approaches in two national communities of communication researches, France and Germany. Radical approaches are nearly invisible in the French community, here social constructivism relies mainly on action and speech theory.”

Education Science


“Drawing on perspectives from a range of different fields (ethics, mathematics education, philosophy, social psychology, science education, social studies), the essays in this book invite us to reposition ourselves in relation to the major currents that have influenced education in this century, namely pragmatism, genetic epistemology, and social interactionism. They call for new reflection on the validity of knowledge and types of knowledge, the compartmentalization of school subjects, the mediating role of teachers, and, above all, the ends of education.”


“This article reports on students’ construction of fraction composition schemes. Students’ whole number multiplicative concepts were found to be critical constructive resources for students’ fraction composition schemes. Specifically, the interiorization of two levels of units, a particular multiplicative concept, was found to be necessary for the construction of a unit fraction composition scheme, while the interiorization of three levels of units was necessary for the construction of a general fraction composition scheme.”


“The purpose of this study is to closely analyze brain-based and constructivist learning approaches, to reveal their relationship, and in light of paradigms and the integral model, to elaborate on the meaning and importance of this relationship for the field of education.”


“Do the ethical aims of research and intervention in education and training not lead us to a ‘new discourse on the study method of our time?’”


“The notion of scientific research programmes is used to understand the development, limitations and potential of constructivism. It is shown that constructivist work in science education fits into a coherent programme exploring the contingencies of learning science. The author goes further to address criticisms of constructivism; evaluate progress in the field; and suggest directions for future research. It is con-
Although constructivism has provided the foundations for a progressive research programme that continues to guide enquiry into learning and teaching science.

**Interdisciplinary**


“This review article argues that ‘situated cognition’ is a manifestation of a broadly shared perspective on the nature of causality in complex systems. Crosscutting disciplinary themes reveal that human cognitive processes are inherently social, interactive, personal, biologically, and neurological, which is to say that a variety of systems develop and depend on one another in complex ways. The concepts, perspectives, and theoretical frameworks that influenced the situated cognition of the 1980s are still alive in potential for thoughtful reconsideration in tomorrow’s cognitive research.”


“Essentialist and constructivist approaches in ethnicity theory are assessed in the case of the Russian constitutional law terminology (ethnic groups designations).”

**Mathematics**


“The authors argue in favor of the view that in mathematics, which increasingly relies on computers to warrant mathematical results, the hunt for absolute certainty will become more and more expensive.”


“The paper gives currency to the humanist case for mathematics as an open-ended endeavor by revealing the irreducible rhetorical nature of mathematics: Mathematical arguments do not stand in isolation of other mathematical arguments. Rather, they form trains of formal and informal arguments, adding up to interconnected theorems, theories and eventually entire fields.”

**Neuroscience**


“According to enactivism, the brain does not process information but rather simulate behaviors. In our experiment, rather than memorizing moving visual targets per se test subjects more globally memorized a situation during which they acted upon the object, i.e., themselves visually tracking the moving target.”


“In our experiment investigating the ‘visuo-motor priming effect’ we show that this effect is due to the interference created by simultaneous unfolding of two cerebral sensorimotor events: manual response motor control and the mental simulation of visual stimulus reach and grasp movement. The data suggest that the main function of the brain is to simulate behaviors.”

**Philosophy**


“This classic French text introduces the epistemology of constructivism and complexity which has been shaped throughout the 20th century. Among other topics, it investigates how ‘valid knowledge’ can be legitimated in the scientific context.”


**Psychology**


“Donald T. Campbell’s evolutionary epistemology is used as a framework for examining five issues facing constructivism: (1) realism, (2) cognitivism, (3) relativism, (4) dualism, and (5) social constructionism. It is suggested that a nuanced application of evolutionary epistemology facilitates fresh ways for constructivists to accommodate these issues.”

http://www2.newpaltz.edu/~raskinj/scholarship/reprints/Evol_of_Constructivism.pdf


“Golden section studies based in Kelly’s personal constructivism”

Raskin J. D. (2009) *Striking the golden section in stigma research*. “Golden section studies based in Kelly’s personal constructivism have a long history within personal construct psychology.”


“This paper examines Kelly’s notion of aggression.”
Guidelines for Authors

STRUCTURE OF THE PAPER

Please make sure that your paper contains the following parts.

- **Title**, optionally subtitle
- **Author(s)** with affiliation(s) and email(s)
- **Meta information**, structured abstract of about 200 words, and up to 6 key words

**Introduction**

The first chapter initializes the contact between author and reader, and should be guided by the question: "Why should the reader get involved with my paper?"

**Main text**

Ideas should be presented in a logical sequence – "Is there a clearly defined progression of information? Does one paragraph lead smoothly into the next?" The writing style should be simple, using as few words as possible. Conciseness and brevity are valued.

**Conclusion**

Provides a summary – "What main points did I make, what did I show?" – and discusses the paper’s relevance – "How is my paper related to constructivist approaches?" Optionally it may provide an outlook – "What could be done next?"

**Alphabetical list of references.** References must not be included as foot-/endnotes

**Biographical sketch and photo of each author.**

- **List of at least 5 potential reviewers.**

LAYOUT & GRAPHICS

- **Overall length:** 3000–9000 words
- **simple single-column format**
- **To emphasize, use italics type** (no bold)
- **As few footnotes as possible**
- **Submission format:** .doc, .rtf, .html
- **For each graphic use a separate file; photographs should have at least 300dpi**

LANGUAGE & COPYRIGHT

- **The paper must be written in English.** If English is a foreign language for you, please ask a native speaker of English to proofread your article before submission
- **The paper must be original work and must not have been published elsewhere**
- **The copyright remains with the author and is licensed under a Creative Commons License, http://creativecommons.org**
- **If you use copyrighted material** (long quotes, photographs, figures, etc.) you must obtain the permission from the respective copyright holder before submitting the final version of your paper.

META INFORMATION

**Paper type:** Which type of inquiry do you follow? Choose from: *conceptual; empirical; synthetic* (formal or computational models); *survey* (guiding summary of a field); *perspective* (of senior researchers)

**Background(s):** Which is the disciplinary background of your paper? Choose from: biological; cognitive; computer science; education science; engineering; epistemological; historical; philosophical; physics; physiological; psychological; sociological

**Perspective:** From which perspective do you argue in your paper? Choose one from: biology of cognition; constructivist evolutionary epistemology; cybernetics; enactive cognitive science; epistemic structuring of experience; non-dualizing philosophy; radical constructivism; second order cybernetics; theory of autopoietic systems.

STRUCTURED ABSTRACT

**Context:** What is the current situation in your discipline with regard to the topic of your paper? Why is it a problem in your discipline at the moment?

**Problem:** What problems do you want to solve? What are the reasons for writing the paper or the aims of the research?

**Method:** What is the approach to the topic and what is the theoretical or subject scope of the paper? How are the objectives achieved? What are the main method(s) used for the research?

**Results:** What was found in the course of the argumentation? What is the solution to the problem you pose?

**Implications:** What is the value of the paper? For whom are your insights beneficial? What do you suggest for future research? Are there identifiable limitations in the research process? What outcomes and implications for practice, applications and consequences are identified? What changes to practice should be made as a result of this paper?

CITING IN THE TEXT

- **Surname of author(s) no comma Year**
- **More than three authors: use the first author’s surname followed by "et al."**
- **All quotes have to be accompanied by a page specification.**
- **Every page specification must be preceded by a colon in both text and reference part.**

**Examples**

- **Langley et al. (1987: 103) showed that…**

**LIST OF REFERENCES**

- **As a rule, use a simplified Harvard-style**
- **Except for the first word paper and book titles are not capitalized**
- **Journal titles are capitalized**
- **No comma between surname and initials**
- **Page specifications are preceded by a colon**
- **Always list all authors (no "et al.")**

**Examples of books**


**Examples of book chapters**


**Examples of journal articles**


**Examples of electronic sources**


**Reprints and translations**

Please cite the reprint or translation from which you quote or which you actually read and add a note about the original publication.

Word template: http://www.univie.ac.at/constructivism/journal/guidelines/template.doc
Table of Contents

117  **Ernst von Glasersfeld**  
Relativism, Fascism, and the Question of Ethics in Constructivism  
*Selected Papers from the Heinz von Foerster Congress 2007 – Part 2*

121  **Louis Kauffman**  
Reflexivity and Eigenform: The Shape of Process

138  **Klaus Krippendorff**  
Conversation: Possibilities of its Repair and Descent into Discourse and Computation

151  **Bernard Scott**  
Conversation, Individuals and Concepts: Some Key Concepts in Gordon Pask’s Interaction of Actors and Conversation Theories

159  **Hans-Herbert Kögler**  
Consciousness as Symbolic Construction: A Semiotics of Thought after Cassirer

170  **Dawid Lubiszewski**  

172  **Przemyslaw Nowakowski**  

174  Publication Review