

## Learning Networks and Connective Knowledge

Stephen Downes  
National Research Council, Canada

### ABSTRACT

The purpose of this chapter is to outline some of the thinking behind new e-learning technology, including e-portfolios and personal learning environments. Part of this thinking is centered around the theory of connectivism, which asserts that knowledge - and therefore the learning of knowledge - is distributive, that is, not located in any given place (and therefore not 'transferred' or 'transacted' per se) but rather consists of the network of connections formed from experience and interactions with a knowing community. And another part of this thinking is centered around the new, and the newly empowered, learner, the member of the net generation, who is thinking and interacting in new ways. These trends combine to form what is sometimes called 'e-learning 2.0' - an approach to learning that is based on conversation and interaction, on sharing, creation and participation, on learning not as a separate activity, but rather, as embedded in meaningful activities such as games or workflows.

**Keywords:** Learning networks, connective knowledge, cognitivism, e-learning 2.0, semantics, ...

### THE TRADITIONAL THEORY: COGNITIVISM

The dominant theory of online and distance learning may be characterized as conforming to a 'cognitivist' theory of knowledge and learning. Cognitivism is probably best thought of as a response to behaviourism. It provides an explicit description of the 'inner workings' of the mind that behaviourism ignores. It is founded on the view that the behaviourist assertion that there are no mental events is in a certain sense implausible, if only by introspection. There is something that it is 'like' to have a belief, and this something seems clearly to be distinct from the mere assemblage of physical constituents. Searle in *Minds, Brains, and Programs* (1980) and Nagel in *What is it Like to Be a Bat* (1974) offer the most compelling versions of this argument.

In other words, cognitivists defend an approach that may be called 'folk psychology' (Ravenscroft, 2004). "In our everyday social interactions we both predict and explain behavior, and our explanations are couched in a mentalistic vocabulary which includes terms like 'belief' and 'desire'." The argument, in a nutshell, is that the claims of folk psychology are literally true, that there is, for example, an entity in the mind corresponding to the belief that 'Paris is the capital of France', and that this belief is, in fact, what might loosely be called 'brain writing' - or, more precisely, there is a one-to-one correspondence between a person's brain states and the sentence itself.

One branch of folk psychology, the language of thought theory, holds that things like beliefs are literally sentences in the brain, and that the materials for such sentences are innate. This is not as absurd as it sounds, and writers like Fodor offer a long and well-argued defense in works such as *The Language of Thought* (2005), *RePresentations* (1983) and *Psychosemantics* (1989). Intuitively, though, you can think of it this way: sculptors sometimes say 'the sculpture was already in the rock; I just found it'. And, quite literally, it makes no sense to say that the sculpture was not in the rock - where else would it be? The idea of 'shaping the mind' is the same sort of thing; it is a revealing of the potential that is latent in the mind, the pre-existing capacity to learn not only language but even sets of concepts and universal truths.

Where the Fodor approach intersects with learning theory is via communication theory, the idea that communication consists of information that flows through a channel (Griffin, 2002). When we join folk psychology with communications theory, we get the idea that there is something like mental content that is in some way transmitted from a sender to a receiver. That we send ideas or beliefs or desires through his channel. Or at the very least, that we send linguistic or non-linguistic (audio music and video images, for example) *representations* of these mental entities.

In learning theory, the concept of transactional distance is based on this sort of analysis of communication (Moore, 1973). What that means is that there exists a space (construed either physically or metaphorically) between two entities between which there exists a channel of communication. In one entity there exists a state, a mental state, which corresponds to a semantic state (in other words, a sentence), and in the process of communication, (aspects of) that state are transmitted from the first entity to the second. This transmission is known as a signal, and as writers like Schramm (1964, 1997) observe, the state transfer is made possible because it constitutes an experience (a mental state) shared between sender and receiver.

This signal, in physical form (such as, say, a book) may constitute an artifact; alternatively, it may be viewed as a medium. The physical analysis of learning, on this account, becomes possible because the physical state - the actual communicative entity - matches the mental state. Thus, the relative states in the sender and the receiver can be (putatively) observed and measured. For example, this approach allows Dretske, in *Knowledge and the Flow of Information* (1999), to explain communication from the perspective of information theory. The transfer of information, suggests Dretske, occurs when, as the result of a signal from an external entity, one's assessment of the total possible states of affairs in the world is reduced.

Moore's (1989, 1993) contribution to educational theory may be placed firmly within this framework. His view is that the effectiveness of communication is improved through interaction. Instead of viewing communication as a one-time event, in which information is sent from a sender and received by a receiver, the transfer of information is enabled through a series of communications, such that the receiver sends messages back to the sender, or to third parties. This is similar to the 'checksum' mechanism in computer communications, where the receiving computer sends back a string of bits to the sender

in order to confirm that the message has been received correctly. Minimally, through this communication, a process of verification is enabled; one can easily infer more complex communications in which knowledge is actually generated, or constructed, by the receiver based on prompts and cues from the sender.

Again, though, notice the pattern here. What is happening is that information theorists, such as Dretske, along with educational theorists, such as Moore, are transferring the properties of a physical medium, in this case, the communication of content via electronic or other signals, to the realm of the mental. Transactional distance just is an application of a physical concept to a mental concept. And if you buy into this, you are bound to buy in to the rest of it, and most especially, that there is something we'll call 'mental content' which is an *isomorphism* between physical states of the brain and the semantical content transmitted to and received by students, who either in some way absorb or construct a mental state that is the same as the teacher's - a 'shared experience'.

### **THE EMERGENTIST ALTERNATIVE AND THE ARGUMENT AGAINST COGNITIVISM**

The allure of a causal theory is also that there appears to be no alternative. If there is no causal connection between teacher and learner, then how can any learning take place, except through some sort of divine intervention? Once we have established and begun to describe the causal process through which information is transacted from teacher to learner, we have pretty much claimed the field; any further account along these lines is an enhancement, an embellishment, but certainly not something new.

There is, however, an alternative. We may contrast cognitivism, which is a *causal* theory of mind, with connectionism, which is an *emergentist* theory of mind. This is not to say that connectionism (Garson, 2007) does away with causation altogether; it is not a 'hand of God' theory. It allows that there is a physical, causal connection between entities, and this is what makes communication possible. But where it differs is, crucially: *the transfer of information does not reduce to this physical substrate*. Contrary to the communications-theoretical account, the new theory is a non-reductive theory. The contents of communications, such as sentences, *are not isomorphic* with some mental state.

Philosophically, there is substantial support for emergentist theories of knowledge (Malpas, 2005). Philosophers have come up with the concept of 'supervenience' to describe something that is not the same as (i.e., not reducible to) physical phenomena, but which are nonetheless dependent on them. Thus, collections of physical states may share the same non-physical state; this non-physical state may be described as a 'pattern', or variously, 'a mental state', 'information', a 'belief', or whatever. Knowledge (and other mental states, concepts, and the like) when represented in this way are 'distributed' - that is, there is no discrete entity that is (or could be) an 'instance' of that knowledge. Computationally, the theory also enjoys support. It is based in one of two major approaches to artificial intelligence. When we think of AI, we usually think of programs and algorithms - the usual stuff of information, signals and channels in computer theory

closely tied to associated concepts in communication theory. The 'General Problem Solver' of Newell and Simon (1963), for example, take a 'symbol processing' approach to computation in AI. This is similar to the Fodor theory, the idea that cognition is (essentially) reducible to a physical symbol set (and therefore instances of cognition and transaction) are governed by the same mechanism. Against this, however, and arguably superior, is the 'connectionist' approach to AI, as described above in the work of Minsky and Papert (1969) or Rumelhart and McClelland (1987).

Mathematically, there is additional support. The properties of networks, as distinct from (typical) causal systems are expressed as a branch of graph theory, the study of which has recently come into prominence because of the work of Watts (2004) and Buchanan (2002). These studies show not only how networks come to be structured as they are but also illustrate how something like, say, a concept can become a 'network phenomenon'. This mathematical description, note these authors, can be used to explain wide varieties of empirical phenomena, from the synchronicity of crickets chirping to the development of trees and river systems.

What grounds this move to networks? On what basis is it proposed that we abandon the traditional conception of learning? In a nutshell, research in mental phenomena has been running in this direction. From early work, such as Marr's *Vision* (1982) and Kosslyn's *Image and Mind* (2002) through Churchland's *Neurophilosophy* (1986) to LeDoux's contemporary *The Synaptic Self* (2002), it is becoming increasingly evident that what we call 'mental contents' do not resemble sentences, much less physical objects, at all.

For example (and there are *many* we could choose from), consider O'Reilly on how the brain represents conceptual structures, as described in *Modeling Integration and Dissociation in Brain and Cognitive Development* (2006). He explicitly rejects the 'isomorphic' view of mental contents, and instead describes a network of distributed representations. "Instead of viewing brain areas as being specialized for specific representational content (e.g., color, shape, location, etc), areas are specialized for specific computational functions by virtue of having different neural parameters... "This 'functionalist' perspective has been instantiated in a number of neural network models of different brain areas, including posterior (perceptual) neocortex, hippocampus, and the prefrontal cortex/basal ganglia system... many aspects of these areas work in the same way (and on the same representational content), and in many respects the system can be considered to function as one big undifferentiated whole. For example, any given memory is encoded in synapses distributed throughout the entire system, and all areas participate in some way in representing most memories." In other words, what O'Reilly is proposing is a functionalist architecture over distributed representation. "*Functionalism* in the philosophy of mind is the doctrine that what makes something a mental state of a particular type does not depend on its internal constitution, but rather on the way it functions, or the role it plays, in the system of which it is a part." For example, when I say, "What makes something a learning object is how we *use* the learning object," I am asserting a functionalist approach to the definition of learning objects (people are so habituated to essentialist definitions that my definition does not even appear on lists of definitions of learning objects). It's like asking, what makes a person a 'bus driver'? Is it

**Comment [CTS1]:** Need paper number or paragraph number here

**Comment [CTS2]:** Need page number or paragraph number here

the colour of his blood? The nature of his muscles? A particular mental state? No - according to functionalism, what makes him a 'bus driver' is the fact that he drives buses. He performs that *function*. "A *distributed representation* is one in which meaning is not captured by a single symbolic unit, but rather arises from the interaction of a set of units, normally in a network of some sort." As noted in the same article, "The concept of distributed representation is a product of joint developments in the neurosciences and in connectionist work on recognition tasks (Churchland and Sejnowski 1992). Fundamentally, a distributed representation is one in which meaning is not captured by a single symbolic unit, but rather arises from the interaction of a set of units, normally in a network of some sort."

**Comment [CTS3]:** Need the page number or paragraph number

**Comment [CTS4]:** Page or paragraph number?

To illustrate this concept, I have been asking people to think of the concept 'Paris'. If 'Paris' were represented by a simple symbol set, we would all mean the same thing when we say 'Paris'. But in fact, we each mean a collection of different things and none of our collections is the same. Therefore, in our own minds, the concept 'Paris' is a loose association of a whole bunch of different things, and hence the concept 'Paris' exists in no particular place in our minds, but rather, is scattered throughout our minds.

Now what the article is saying is that human brains *are* like computers - but not like the computers as described above, with symbols and programs and all that, but like computers when they are connected together in a network. "The brain as a whole operates more like a social network than a digital computer... the computer-like features of the prefrontal cortex broaden the social networks, helping the brain become more flexible in processing novel and symbolic information." Understanding 'where the car is parked' is like understanding how one kind of function applies on the brain's distributed representation, while understanding 'the best place to park the car' is like how a *different* function applies to the *same* distributed representation.

**Comment [CTS5]:** Page or paragraph number citation?

The analogy with the network of computers is a good one (and people who develop social network software are sometimes operating with these concepts of neural mechanisms specifically in mind). The actual social network itself - a set of distributed and interlinked entities, usually people, as represented by websites or pages - constitutes a type of distributed representation. A 'meme' - like, say, the *Friday Five* - is distributed across that network; it exists in no particular place.

Specific mental operations, therefore, are like thinking of functions applied to this social network. For example, if I were to want to find 'the most popular bloggers' I would need to apply a set of functions to that network. I would need to represent each entity as a 'linking' entity. I would need to cluster types of links (to eliminate self-referential links and spam). I would then need to apply my function (now my *own* view here, and possibly O'Reilly's, though I don't read it specifically in his article, is that to apply a function is to create *additional neural layers* that act as specialized filters - this would contrast with, say, Technorati, which polls each individual entity and then applies an *algorithm* to it).

This theory, stated simply, is that human thought amounts to patterns of interactions in neural networks. More precisely, patterns of input phenomena - such as sensory

perceptions - cause or create patterns of connections between neurons in the brain. These connections are associative - that is, connections between two neurons form when the two neurons are active at the same time, and weaken when they are inactive or active at different times. See, for example, Hebb's *The Organization of Behavior* (2002), which outlines what has come to be called 'Hebbian associationism'.

## THE ARGUMENT AGAINST COGNITIVISM

As we examine the emergentist theory of mind we can arrive at five major implications of this approach for educational theorists:

- first, knowledge is subsymbolic. Mere possession of the words does not mean that there is knowledge; the possession of knowledge does not necessarily result in the possession of the words (and for much more on this, see Polanyi's (1974) discussion of 'tacit knowledge' in 'Personal Knowledge').
- second, knowledge is distributed. There is no specific 'mental entity' that corresponds to the belief that 'Paris is the capital of France'. What we call that 'knowledge' is (an indistinguishable) pattern of connections between neurons. See, for example, Geoffrey Hinton(1986), *Learning Distributed Representations of Concepts*.
- third, knowledge is interconnected. The same neuron that is a part of 'Paris is the capital of France' might also be a part of 'My dog is named Fred'. It is important to note that this is a non-symbolic interconnection - this is the basis for non-rational associations, such as are described in the article, *Where Belief is Born* (Jha, 2005)
- fourth, knowledge is personal. Your 'belief' that 'Paris is the capital of France' is quite literally different from my belief that 'Paris is the capital of France'. If you think about it, this must be the case - otherwise Gestalt tests would be useless; we would all utter the same word when shown the same picture.
- fifth, what we call 'knowledge' (or 'belief', or 'memory') is an emergent phenomenon. Specifically, it is not 'in' the brain itself, or even 'in' the connections themselves, because there is no 'canonical' set of connections that corresponds with 'Paris is the capital of France'. It is, rather (and carefully stated), a recognition of a pattern in a set of neural events (if we are introspecting) or behavioural events (if we are observing). We infer to mental contents the same way we watch Donald Duck on TV - we think we see something, but that something is not actually there - it's just an organization of pixels.

This set of features constitutes a mechanism for evaluating whether a cognitivist theory or a connectivist theory is likely to be true. In my own mind (and in my own writing, as this was the subject of my first published paper, 'Why Equi Fails'), the mechanism can be summed in one empirical test: context sensitivity. If learning is context-sensitive then the 'language of thought' hypothesis fails, and the rest of folk psychology along with it. For the presumption of these theories is that, when you believe that 'Paris is the capital of France' and when I believe that 'Paris is the capital of France', that we believe the *same thing*, and that, importantly, *we share the same mental state*, and hence can be reasonably

relied upon to demonstrate the same semantic information when prompted.

So I've concluded that 'language of thought' hypothesis could not possibly succeed, nor folk psychology either. Because it turns out that not only language but the whole range of phenomena associated with knowledge and learning are context-sensitive. Or so the philosophers say.

- Wittgenstein, in *Philosophical Investigations* (1973) and elsewhere, argues that meaning is context sensitive, that what we mean by a word depends on a community of speakers; there is no such thing as a 'private language', and hence, the meaning of a word cannot stand alone, fully self-contained, in the mind.
- Quine, in *Two Dogmas of Empiricism* (1951) and in *Word and Object* (1964), shows that observation itself is context-sensitive, that there is no knowable one-to-one matching between sense-phenomena and the words used to describe them; in 'On the Indeterminacy of Translation' he illustrates this with the famous 'gavagai' example: when a native speaker uses the word 'gavagai' there is no empirical way to know whether he means 'rabbit' or 'the physically incarnate manifestation of my ancestor'
- Hanson, in *Patterns of Discovery* (1958), argues, in my view successfully, that causal explanations are context-sensitive. 'What was the cause of the accident?' It depends on who you ask - the police officer will point to the speed, the urban planner will point to the road design, the driver will point to the visibility.
- Lakoff, in *Women, Fire and Dangerous Things* (1990), shows that categories are context sensitive (contra Saul Kripke); that what makes two things 'the same' varies from culture to culture, and indeed (as evidenced from some of his more recent political writings) from 'frame' to 'frame'.
- van Fraassen in *The Scientific Image* (1980) shows that explanations are context sensitive. 'Why are the roses growing here?' may be answered in a number of ways, depending on what alternative explanations are anticipated. 'Because someone planted them.' 'Because they were well fertilized.' 'Because the chlorophyll in the leaves converts the energy of the Sun into glucose' are all acceptable answers, the correct one of which depends on the presuppositions inherent in the question.
- Lewis (2001) and Stalnaker (1987) argue that the counterfactuals and modalities are context sensitive (though Lewis, if asked, would probably deny it). The truth of a sentence like 'brakeless trains are dangerous' depends, not on observation, but rather, on the construction of a 'possible world' that is relevantly similar (Stalnaker uses the word 'salience') to our own, but what counts as 'relevant' depends on the context in which the hypothetical is being considered.

If, as asserted above, what counts as knowledge of even basic things like the meanings of words and the cause of events is sensitive to context, then it seems clear that such knowledge is not a stand-alone symbolic *representation* of that knowledge, since representations would not be, could not be, context sensitive. Rather, what is happening is that each person is experiencing a mental state that is at best seen as an *approximation* of what it is that is being said in words or experienced in nature, an approximation that is

framed and indeed comprehensible only from which the rich set of world views, previous experiences and frames in which it embedded.

If this is the case, then the concepts of *what it is to know* and *what it is to teach* are very different from the traditional theories that dominate distance education today. Because if learning is not the transfer of mental contents – if there is, indeed, no such mental content that exists to be transported – then we need to ask, what is it that we are attempting to do when we attempt to teach and learn.

### **NETWORK SEMANTICS AND CONNECTIVE LEARNING**

If we accept that something like the network theory of learning is true, then we are faced with a knowledge and learning environment very different from what we are used to. In the strictest sense, there is no semantics in network learning, because there is no *meaning* in network learning (and hence, the constructivist practice of ‘making meaning’ is literally meaningless).

Traditionally, what a sentence ‘means’ is the (truth or falsity of) the state of the world it represents. However, on a network theory of knowledge, there is no such state of the world to which this meaning can be affixed. This is not because there is no such state of the world. The world could most certainly exist, and there is no contradiction in saying that a person’s neural states are caused by world events. However, it *does* mean that there is no *particular* state of the world that corresponds with (is isomorphic to) a particular mental state. This is because the mental state is embedded in a sea of context and presuppositions that are completely opaque to the state of the world.

How, then, do we express ourselves? How do we distinguish between true and false – what, indeed, does it even *mean* to say that something is true and false? The answer to these questions is going to be different for each of us. They will be embedded in a network of assumptions and beliefs about the nature of meaning, truth and falsity. In order to get at a response, therefore, it will be necessary to outline what may only loosely be called ‘network semantics’.

We begin with the nature of a network itself. In any network, there will be three major elements:

- Entities, that is, the things that are connected that send and receive signals
- Connections, that is, the link or channel between entities (may be represented as physical or virtual)
- Signals, that is, the message sent between entities. Note that meaning is *not inherent* in signal and must be interpreted by the receiver

In an environment of this description, then, networks may vary according to a certain set of properties:

- Density, or how many other entities each entity is connected to

- Speed, or how quickly a message moves to an entity (can be measured in time or 'hops')
- Flow, or how much information an entity processes, which includes messages sent and received in addition to transfers of messages for other entities
- Plasticity, or, how frequently connections created, abandoned
- Degree of connectedness – is a function of density, speed, flow and plasticity

Given this description of networks, we can identify the essential elements of network semantics.

- First, *context*, that is, the localization of entities in a network. Each context is unique – entities see the network differently, experience the world differently. Context is required in order to interpret signals, that is, each signal means something different depending on the perspective of the entity receiving it.
- Second, *salience*, that is, the relevance or importance of a message. This amounts to the similarity between one pattern of connectivity and another. If a signal creates the activation of a set of connections that were previously activated, then this signal is salient. Meaning is created from context and messages via salience.
- Third, *emergence*, that is, the development of patterns in the network. Emergence is a process of resonance or synchronicity, not creation. We do not *create* emergent phenomena. Rather emergence phenomena are more like commonalities in patterns of perception. It requires an interpretation to be recognized; this happens when a pattern becomes *salient* to a perceiver.
- Fourth, *memory* is the persistence of patterns of connectivity, and in particular, those patterns of connectivity that result from, and result in, salient signals or perceptions.

Given this background, what does it mean, then, to say that a sentence has semantical import? To say, similarly, that we 'know' something? As suggested above, most of us remain committed to something like a traditional (Tarski) semantics: we know something just in case what we know happens to be true. But of course, this fails to tell the whole story. The knowledge needs to be, in some way, in our mind (or in our society); it needs to be a 'belief'. And (so goes the argument) it needs to be in some way justified, through a process of verification, or at the very least, says Popper, through the absence of falsification.

Significant difficulties emerge when we try to articulate what it is that we know. Consider, for example, 'snow is white'. Sure, one could check some snow in order to determine that it is white, but only if one first understood what is meant by 'snow' and 'white' (not to mention, as Clinton taught us, 'is'). But as discussed above, that constitutes *the* meaning of, ay, 'snow', is far from clear. There is no such single entity. What it means is a matter of interpretation. So, for example, does enumerating what constitutes instance of snow. Does 'yellow snow' count? Does snow produced by artificial ice machines count?

From the discussion above, it should be clear that on the account being given here, to 'know' that 'snow is white' is to be *organized in a certain way* (one that is evidenced by uttering 'snow' when asked). To be organized in such a way as to have neural and mental structures corresponding to the words 'snow', 'is' and 'white', where those structures are such that the concept 'snow' is closely associated with (in certain contexts) the concept 'white' (obviously this is a gloss, since there is no real correspondence). Knowing that 'snow is white' is therefore being organized in *some* certain way, but not in a *specific* particular way (we couldn't examine one's neural organization and be able to say whether the person knows that snow is white).

What it means to 'know' then is based on organization and connectedness in the brain. What it is to 'know' is, if you will, a natural development that occurs in the mind when it is presented with certain sets of phenomena; other things being equal, present the learner with different phenomena and they will learn different things.

Whether something counts as 'knowledge' rather than, say, 'belief' or 'speculation', depends less on the state of the world, and more on the strength or degree of connectedness between the entities. To 'know' something is to not be able to not know. It's like finding Waldo, or looking at an abstract image. There may be a time when we don't know where Waldo is, or what the image represents, but once we have an interpretation, it is not possible to look without seeing Waldo, without seeing the image. No wonder Dreyfus and Dreyfus (1986) talk about 'levels' of knowledge, up to and including an almost intuitive 'expert' knowledge. As a particular organization, a particular set of connections, between neural structures is strengthened, as this structure becomes embedded in more and more of our other concepts and other knowledge, it changes its nature, changing from something that needs to be triggered by cue or association (or mental effort) into something that is natural as other things we 'know' deeply, like how to breathe, and how to walk, structures entrenched through years, decades, or successful practice. Contrast this to a cognitivist model of knowledge, where once justification is presented, something is 'known', and cannot later in life be 'more known'.

Connective semantics is therefore derived from what might be called connectivist 'pragmatics', that is, that actual *use* of networks in practice. In our particular circumstance we would examine how networks are used to support learning. The methodology employed is to look at multiple examples and to determine what patterns may be discerned. These patterns cannot be directly communicated. But instances of these patterns may be communicated, thus allowing readers to (more or less) 'get the idea'.

For example, in order to illustrate the observation that 'knowledge is distributed' I have frequently appealed to the story of the 747. In a nutshell, I ask, "who knows how to make a 747 fly from London to Toronto?" The short answer is that *nobody* knows how to do this – no one person could design a 747, manufacture the parts (including tires and aircraft engines), take it off, fly it properly, tend to the passengers, navigate, and land it successfully. The knowledge is *distributed* across a network of people, and the

phenomenon of ‘flying a 747’ can exist at all only because of the connections between the constituent members of that network.

Or, another story: if knowledge is a network phenomenon, then, is it necessary for all the elements of a bit of knowledge to be stored in one’s own mind? Karen Stephenson (1998) writes, “I store my knowledge in my friends.” This assertion constitutes an explicit recognition that what we ‘know’ is embedded in our network of connections to each other, to resources, to the world. Siemens (2004) writes, “Self-organization on a personal level is a micro-process of the larger self-organizing knowledge constructs created within corporate or institutional environments. The capacity to form connections between sources of information, and thereby create useful information patterns, is required to learn in our knowledge economy.”

This approach to learning has been captured under the heading of ‘connectivism’. In his paper of the same name, George Siemens articulates the major theses:

- Learning and knowledge rests in diversity of opinions.
- Learning is a process of connecting specialized nodes or information sources.
- Learning may reside in non-human appliances.
- Capacity to know more is more critical than what is currently known
- Nurturing and maintaining connections is needed to facilitate continual learning.
- Ability to see connections between fields, ideas, and concepts is a core skill.
- Currency (accurate, up-to-date knowledge) is the intent of all connectivist learning activities.
- Decision-making is in itself a learning process. Choosing what to learn and the meaning of incoming information is seen through the lens of a shifting reality. While there is a right answer now, it may be wrong tomorrow due to alterations in the information climate affecting the decision.

Is this the definitive statement of network learning? Probably not. But it is developed in the classic mold of network learning, through a process of immersion into the network and recognition of salient patterns. What sort of network? The following list is typical of what might be called ‘network’ practices online:

Practice: Content Authoring and Delivery

- Numerous content authoring systems on the web...
- Weblogs – Blogger, Wordpress, LiveJournal, Moveable Type, more
- Content Management Systems – Drupal, PostNuke, Plone, Scoop, and many more...
- Audio – Audacity – and audioblogs.com – and Podcasting
- Digital imagery and video – and let’s not forget Flickr
- Collaborative authoring – Writely, Hula, the wiki

Practice: Organize, Syndicate Sequence, Deliver

- Aggregation of content metadata – RSS and Atom, OPML, FOAF, even DC and LOM

- Aggregators – NewsGator, Bloglines – Edu\_RSS
- Aggregation services – Technorati, Blogdex, PubSub
- More coming – the Semantic Social Network

Practice: Identity and Authorization

- A raft of centralized (or Federated) approaches – from Microsoft Passport to Liberty to Shibboleth
- Also various locking and encryption systems
- But nobody wants these
- Distributed DRM – Creative Commons, ODRL...
- Distributed Identification management – Sxip, LID...

Practice: Chatting, Phoning, Conferencing

- Bulletin board systems and chat rooms, usually attached to the aforementioned content management systems such as Drupal, Plone, PostNuke, Scoop
- Your students use this, even if you don't: ICQ, AIM, YIM, and some even use MSN Messenger
- Audioconferencing? Skype...Or NetworkEducationWare...
- Videoconferencing? Built into AIM... and Skype

## THE MOVE TO 2.0

It is now ten years or so into the era of online learning. Schools, colleges and universities have now developed the internet infrastructure of their choice. Almost all have web pages, most have online courses, and many have synchronous online learning. The learning management system (LMS) has become a commodity business, educational software of all sorts abounds, and the phenomenon has spread around the globe.

Even so, it may be observed that most people online of school or college age are elsewhere. They may not be writing class essays, but they are writing blogs, perhaps one of the 50 million or more tracked by Technorati (2006). They are at MySpace, which counts some 86 million accounts (Kirkpatrick, 2006). They are recording videos, making YouTube even larger than MySpace. They are, in fact, engaged in the many networking activities described in the previous section. *Something* is going on.

On the web, what has happened has been described as the migration to something called *Web 2.0* (pronounced 'web two point oh'). The term, popularized by publisher O'Reilly (2005), describes the evolution of the web into the 'read-write' web. O'Reilly writes, "The central principle behind the success of the giants born in the Web 1.0 era who have survived to lead the Web 2.0 era appears to be this, that they have embraced the power of the web to harness collective intelligence."

As an example, he cites the difference between Netscape and Google. According to O'Reilly, Netscape saw the web as a software market. By releasing its popular browser for free and hence effectively controlling web standards, the company could gain a lock on the web server software market. Google, by contrast, never viewed the web as a place

to ship product. Rather, it became a service, harnessing the collective linking behaviour of web users to create a more effective search engine.

The term 'Web 2.0', as has been widely noted, is a notoriously fuzzy term, difficult to nail down. O'Reilly (2005) offered one set of criteria to describe the difference:

---

Web 1.0	Web 2.0
	DoubleClick --> Google AdSense
	Ofoto --> Flickr
	Akamai --> BitTorrent
	mp3.com --> Napster
	Britannica Online --> Wikipedia
	personal websites --> blogging
	evite --> upcoming.org and EVDB
	domain name speculation --> search engine optimization
	page views --> cost per click
	screen scraping --> web services
	publishing --> participation
	content management systems --> wikis
	directories (taxonomy) --> tagging ("folksonomy")
	stickiness --> syndication

---

This list is incomplete, of course, and as with any definition by example, ultimately unsatisfactory. Nonetheless, the definition may be characteristic of Web 2.0. "That the term has enjoyed such a constant morphing of meaning and interpretation is, in many ways, the clearest sign of its usefulness. This is the nature of the conceptual beast in the digital age, and one of the most telling examples of what Web 2.0 applications do: They replace the authoritative heft of traditional institutions with the surging wisdom of crowds." (Madden & Fox, 2006, p.2).

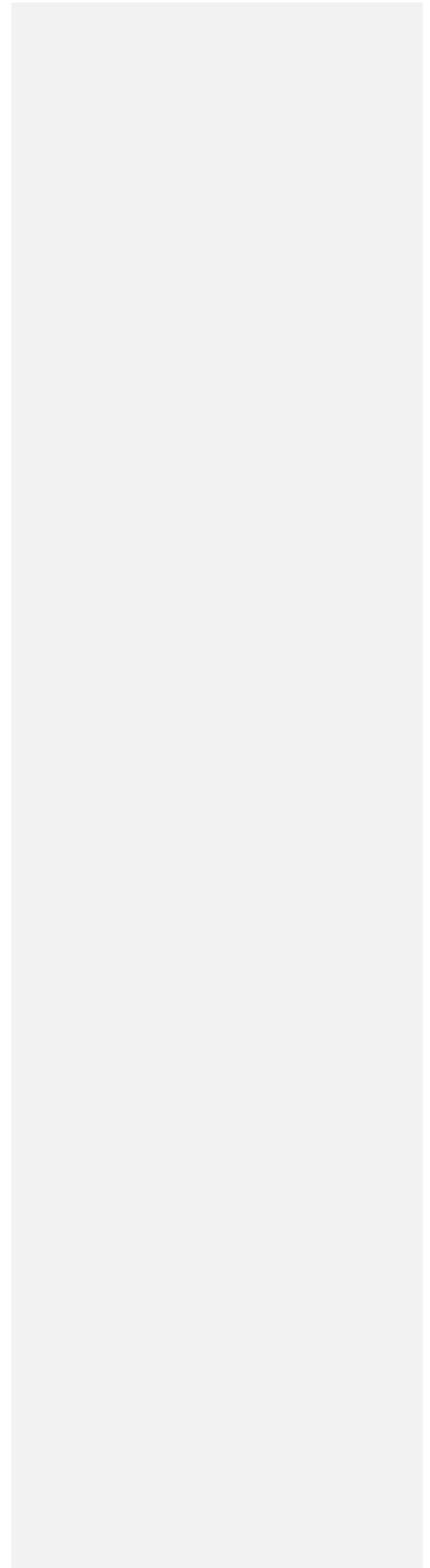
As the web surged toward 2.0 the educational community solidified its hold on the more traditional approach. The learning management system became central (and centralized, with Blackboard purchasing WebCT). Developers continue to emphasize content and software development, as Learning Object Metadata was standardized and IMS developed specifications for content packaging and learning design.

Even so, as traditional instructional software became entrenched, it became difficult not to notice the movement in the other direction. First was the exodus from commercial software in favour of open source systems such as Moodle, Sakai and LAMS. Others eschewed educational software altogether as a wave of educators began to look at the use of blogging and the wiki in their classes. A new, distributed, model of learning was emerging, which came to be characterized as e-learning 2.0 (Downes, 2005).

"What happens," I asked, "when online learning ceases to be like a medium, and becomes more like a platform? What happens when online learning software ceases to be

a type of content-consumption tool, where learning is "delivered," and becomes more like a content-authoring tool, where learning is created?" The answer turns out to be a lot like Web 2.0: "The model of e-learning as being a type of content, produced by publishers, organized and structured into courses, and consumed by students, is turned on its head. Insofar as there is content, it is used rather than read— and is, in any case, more likely to be produced by students than courseware authors. And insofar as there is structure, it is more likely to resemble a language or a conversation rather than a book or a manual."

In the days since this shift was recognized a growing community of educators and developers has been gathering around a model of online learning typified by this diagram (Figure 1) authored by Wilson et al. (2006), and remixed by various others since then.



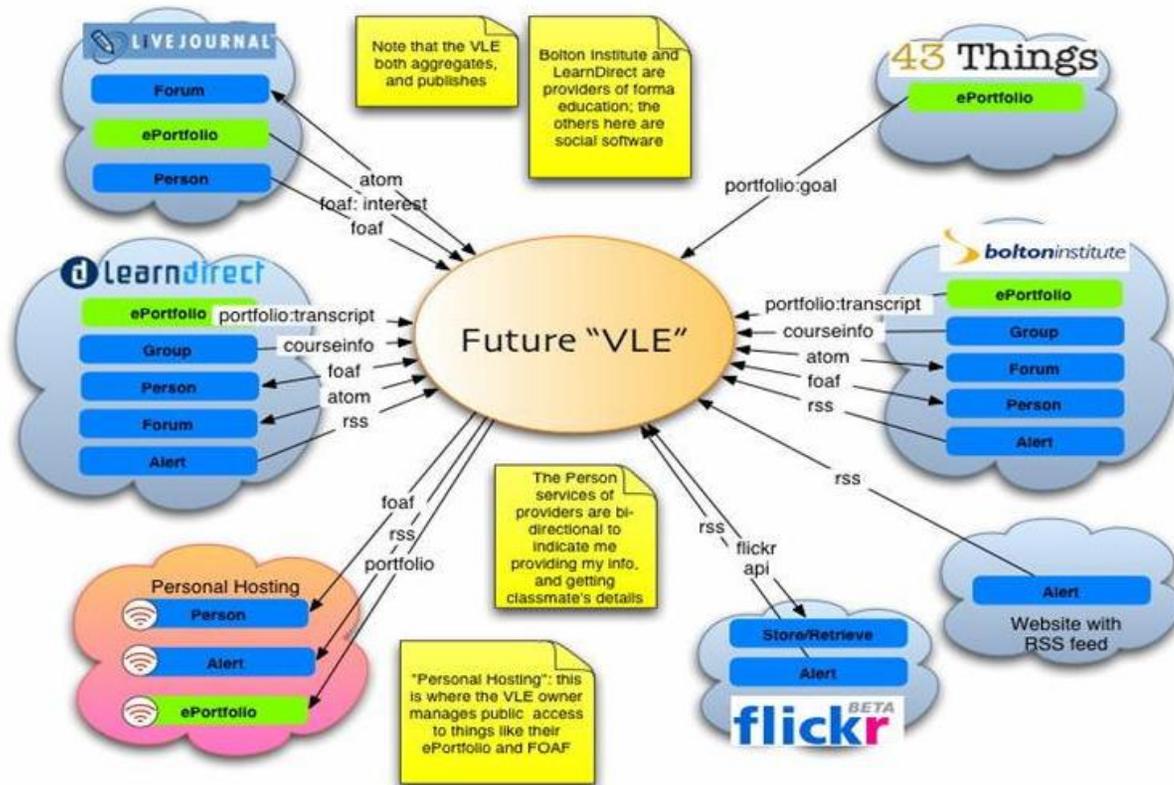


Figure 1: Future VLE

<http://www.cetis.ac.uk/members/scott/blogview?entry=20050125170206>

The 'future VLE' is now most commonly referred to as the 'Personal Learning Environment', or PLE. As described by Milligan (2006), PLEs "would give the learner greater control over their learning experience (managing their resources, the work they have produced, the activities they participate in) and would constitute their own personal learning environment, which they could use to interact with institutional systems to access content, assessment, libraries and the like." (¶ 4).

The idea behind the personal learning environment is that the management of learning migrates from the institution to the learner. As the diagram shows, the PLE connects to a number of remote services, some that specialize in learning and some that do not. Access to learning becomes access to the resources and services offered by these remote services. The PLE allows the learner not only to consume learning resources, but to produce them as well. Learning therefore evolves from being a transfer of content and knowledge to the production of content and knowledge.

E-learning 2.0 promises a lot. "Like the web itself, the early promise of e-learning - that of empowerment - has not been fully realized. The experience of e-learning for many has been no more than a hand-out published online, coupled with a simple multiple-choice quiz. Hardly inspiring, let alone empowering. But by using these new web services, e-learning has the potential to become far more personal, social and flexible." These technologies, in other words, would *empower* students in a way previous technologies didn't.

But the structure seems to deliver on the promise. As O'Hear (2006) writes, "The traditional approach to e-learning... tends to be structured around courses, timetables, and testing. That is an approach that is too often driven by the needs of the institution rather than the individual learner. In contrast, e-learning 2.0 takes a 'small pieces, loosely joined' approach that combines the use of discrete but complementary tools and web services - such as blogs, wikis, and other social software - to support the creation of ad-hoc learning communities." (¶ 2).

## THE 2.0 ARCHITECTURE

The idea of e-learning 2.0 may appear elusive at first blush, but many of the ideas central to e-learning 2.0 may be evoked through a discussion of its fundamental architecture, which may be called 'learning networks'. The objective of a theory of learning networks is to describe the manner in which resources and services are organized in order to offer learning opportunities in a network environment. Learning networks is not therefore a pedagogical principle, but rather, a description of an environment intended to support a particular pedagogy.

I introduced learning networks formally in my Buntine Oration (2004): "If, as I suggested above, we describe learning objects using the metaphor of language, text, sentences and books, then the metaphor to describe the learning network as I've just described it is the ecosystem, a collection of different entities related in a single

environment that interact with each other in a complex network of affordances and dependencies, an environment where the individual entities are not joined or sequenced or packaged in any way, but rather, live, if you will, free, their nature defined as much by their interactions with each other as by any inherent property in themselves. “We don’t present these learning objects, ordered, in a sequence, we present randomly, unordered. We don’t present them in classrooms and schools, we present them to the environment, to where students find themselves, in their homes and in their workplaces. We don’t present them at all, we contribute them to the conversation, and we become part of the conversation. They are not just text and tests; they are ourselves, our blog posts, our publications and speeches, our thoughts in real-time conversation. Sigmund Freud leaning on the lamp post, just when we need him.”

**Comment [CTS6]:** Citation? Need page number or paragraph number.

This ‘ecosystem’ approach, realized in software, is based on a ‘distributed’ model of resources, as suggested by the PLE diagram. The difference between the traditional and decentralized approach may be observed in the following diagram (Figure 2):

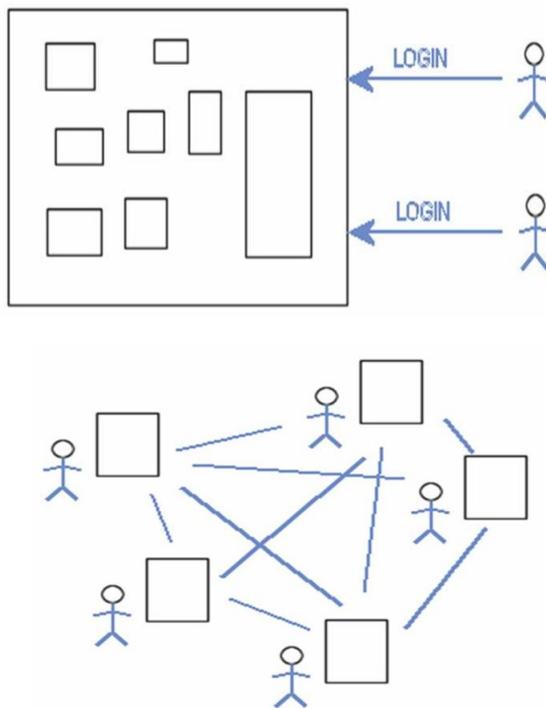


Figure 2: Centralized approach (above) and distributed approach (below)

It is interesting, and worth noting, that before the World Wide Web burst onto the scene, online access in general was typified by the centralized approach depicted in the upper figure. Users would dial up and log on to services such as CompuServe and Prodigy.

The World Wide Web, by contrast, is an example of a distributed environment. There is no single big server; resources and access are scattered around the world in the form of a network of connected web servers and internet service providers. Users do not log into a single service called 'The Web' but are also distributed, accessing through internet service providers. Even their software is distributed; their web browsers run locally, on their own machine, and function by connecting to online services and resources.

In an environment such as this, the nature of design changes. In a typical computer program, the design will be specified with an algorithm or flowchart. Software will be described as performing a specific process, with specified (and often controlled) inputs and outputs. In a distributed environment, however, the design is no longer defined as a type of process. Rather, designers need to characterize the nature of the *connections* (Kraan, 2005) between the constituent entities.

What are the core principles (Kilkki, 2005) that will characterize such a description? The internet itself illustrates a sound set of principles, grounded by two major characteristics: simple services with realistic scope. "Simple service or simple devices with realistic scope are usually able to offer a superior user experience compared to a complex, multi-purpose service or device." Or as Weinberger (2002) describes the network: small pieces, loosely joined.

In practice, these principles may be realized in the following design principles. It is worth noting at this juncture that these principles are intended to describe not only networks but also network learning, to show how network learning differs from traditional learning.

The idea is that each principle confers an advantage over non-network systems, and that the set, therefore, may be used as a means of evaluating new technology. This is a *tentative* set of principles, based on observation and pattern recognition. It is not a definitive list, and indeed, it is likely that there cannot be a definitive list.

1. Effective networks are *decentralized*. Centralized networks have a characteristic 'star' shape, where some entities have many connections while the vast majority have few. This is typical of, say a broadcast network or the method of a teacher in a classroom. Decentralized networks, by contrast, form a mesh. The weight of connections and the flow of information is distributed. This balanced load results in a more stable network, with no single point of failure.
2. Effective networks are *distributed*. Network entities reside in different physical locations. This reduces the risk of network failure. It also reduces need for major infrastructure, such as powerful servers, large bandwidth, and massive storage. Examples of distributed networks include peer-to-peer networks, such as Kazaa, Gnutella and content syndication networks, such as RSS. The emphasis of such systems is on sharing, not copying; local copies, if they exist, are temporary.
3. Effective networks *disintermediated*. That is, they eliminate 'mediation', the barrier between source and receiver. Examples of disintermediation include the

bypassing of editors, replacing peer review prior to publication with recommender systems subsequent to publication. Or of the replacement of traditional news media and broadcasters with networks of news bloggers. And, crucially, the removal of the intermediate teacher that stands between knowledge and the student. The idea is to, where possible, provide direct access to information and services. The purpose of mediation, if any, is to manage flow, not information, to reduce the volume of information, not the type of information.

4. In effective networks, content and services are *disaggregated*. Units of content should be as small as possible and content should not be 'bundled'. Instead, the organization and structure of content and services is created by the receiver. This allows the integration of new information and services with the old, of popular news and services with those in an individual's particular niche interests. This was the idea behind learning objects; the learning object was sometimes defined as the 'smallest possible unit of instruction' (Bath, 2003). The assembly of learning objects into pre-packaged 'courses' defeats this, however, obviating any advantage the disaggregating of content may have provided.
5. In an effective network, content and services are *dis-integrated*. That is to say, entities in a network are not 'components' of one another. For example, plug-ins or required software to be avoided. What this means in practice is that the structure of the message is logically distinct from the type of entity sending or receiving it. The message is coded in a common 'language' where the code is open, not proprietary. So no particular software or device is needed to receive the code. This is the idea of standards, but where standards evolve rather than being created, and where they are adopted by agreement, not requirement.
6. An effective network is *democratic*. Entities in a network are autonomous; they have the freedom to negotiate connections with other entities, and they have the freedom to send and receive information. Diversity in a network is an asset, as it confers flexibility and adaptation. It also allows the network as a whole to represent more than just the part. Control of the entities in a network, therefore, should be impossible. Indeed, in an effective network, even where control seems desirable, it is not practical. This condition – which may be thought of as the *semantic* condition – is what distinguishes networks from groups (see below).
7. An effective network is *dynamic*. A network is a fluid, changing entity, because without change, growth and adaptation are not possible. This is sometimes described as the 'plasticity' of a network. It is through this process of change that new knowledge is discovered, where the creation of connections is a core function.
8. An effective network is *desegregated*. For example, in network learning, learning is not thought of as a Separate Domain. Hence, there is no need for learning-specific tools and processes. Learning is instead thought of as a part of living, of work, of play. The same tools we use to perform day-to-day activities are the tools we use to learn. Viewed more broadly, this condition amounts to seeing the network as infrastructure. Computing, communicating and learning are not something we 'go some place to do'. Instead, we think of network resources as similar to a utility, like electricity, like water, like telephones. The network is everywhere.

It should be noted that though some indication of the justification for these methodological principles has been offered in the list above, along with some examples, this list is in essence *descriptive*. In other words, what is claimed here is that successful networks *in fact* adhere to these principles. The *why* of this is the subject of the next few sections.

## THE SEMANTIC CONDITION

Knowledge is a network phenomenon. To 'know' something is to be organized in a certain way, to exhibit patterns of connectivity. To 'learn' is to acquire certain patterns. This is as true for a community as it is for an individual. But it should be self-evident that mere organization is not the *only* determinate of what constitutes, if you will, 'good' knowledge as opposed to 'bad' (or 'false') knowledge. Consider public knowledge. People form themselves into communities, develop common language and social bonds, and then proceed to invade Europe or commit mass suicide. Nor is personal knowledge any reliable counterbalance to this. People are as inclined to internalize the dysfunctional as the utile, the self-destructive as the empowering. Some types of knowledge (that is, some ways of being organized, whether socially or personally) are destructive and unstable. These are examples of cascade phenomena. In social sciences the same phenomenon might be referred to as the bandwagon effect. Such phenomena exist in the natural world as well. The sweep of the plague through medieval society, the failure of one hydro plant after another, the bubbles in the stock market. Cascade phenomena occur when some event or property sweeps through the network. Cascade phenomena are in one sense difficult to explain and in another sense deceptively simple.

The sense in which they are simple to explain is mathematical. If a signal has more than an even chance of being propagated from one entity in the network to the next, and if the network is fully connected, then the signal will eventually propagate to every entity in the network. The speed at which this process occurs is a property of the connectivity of the network. In (certain) random and scale free networks, including hierarchal networks, it takes very few connections to jump from one side of the network to the other. Cascade phenomena sweep through densely connected networks very rapidly.

The sense in which they are hard to explain is related to the question of why they exist at all. Given the destructive nature of cascade phenomena, it would make more sense to leave entities in the network unconnected (much like Newton escaped the plague by isolating himself). Terminating all the connections would prevent cascade phenomena. However, it would also prevent any possibility of human knowledge, any possibility of a knowing society.

It is tempting to suppose that we could easily surmount the excesses of cascading communities through a simple application of knowledge obtained through other domains, but in practice we gain no increased certainty or security. Nothing guarantees truth. We are as apt to be misled by the information given by our senses, for example, as by any wayward community. Descartes records simple examples, such as mirages, or the bending of a stick in water, to make the point. Today's science can point to much deeper scepticism.

Perception itself consists of selective filtering and interpretation (pattern detection!). The mind supplies sensations that are not there. Even a cautiously aware and reflective perceiver can be misled.

Quantitative knowledge, the cathedral of the twentieth century, fares no better. Though errors in counting are rare, it is a fragile a process. *What* we count is as important as how we count, and on this, quantitative reasoning is silent. We can measure grades, but are grades the measure of learning? We can measure economic growth, but is an increase in the circulation of money a measure of progress? We can easily mislead ourselves with statistics, as Huff (1993) shows, and in more esoteric realms, such as probability, our intuitions can be exactly wrong.

We compensate for these weaknesses by recognizing that a single point of view is insufficient; we distribute what constitutes an 'observation' through a process of description and verification. If one person says he saw a zombie, we take such a claim sceptically; if a hundred people say they saw zombies, we take it more seriously, and if a process is described whereby anyone who is interested can see a zombie for themselves, the observation is accepted. In other words, the veracity of our observations is not guaranteed by the observation, but by an observational *methodology*.

In quantitative reasoning, we take care to ensure that, in our measurements, we are measuring the same thing. Through processes such as double-blind experimentation, we additionally take care to ensure that our expectations do not influence the count. In statistical reasoning, we take care to ensure that we have a sufficiently random and representative sample, in order to ensure that we are measuring one phenomenon, and not a different, unexpected phenomenon. In both we employ what Carnap called the requirement of the total evidence: we peer at something from all angles, all viewpoints, and if everybody (or the preponderance of observers) conclude that it's a duck, then it's a duck.

Connective knowledge is supported through similar mechanisms. It is important to recognize that a structure of connections is, at its heart, *artificial*, an *interpretation* of any reality there may be, and moreover, that our observations of emergent phenomena themselves as fragile and questionable as observations and measurements - these days, maybe more so, because we do not have a sound science of network semantics.

In a network, a cascade phenomenon is akin to jumping to a conclusion about an observation. It is, in a sense, a rash and unthinking response to whatever phenomenon prompted it. This capacity is crucially dependent on the structure of the network. Just as a network with no connections has no capacity to generate knowledge, a fully connected network has no defense against jumping to conclusions. What is needed is to attain a middle point, where full connectivity is achieved, but where impulses in the network ebb and flow, where impulses generated by phenomena are checked against not one but a multitude of competing and even contradictory impulses.

This is what the human mind does naturally. It is constructed in such a way that no single impulse is able to overwhelm the network. A perception must be filtered through layers

of intermediate (and (anthropomorphically) sceptical) neurons before forming a part of a concept. For every organization of neurons that achieves an active state, there are countless alternative organizations ready to be activated by the same, or slightly different, phenomena (think of how even a seed of doubt can destabilize your certainty about something).

Knowledge in the mind is not a matter of mere numbers of neurons being activated by a certain phenomenon; it is an ocean of competing and conflicting possible organizations, each ebbing and subsiding with any new input (or even upon reflection). In such a diverse and demanding environment only patterns of organization genuinely successful in some important manner achieve salience, and even fewer become so important we cannot let them go. In order therefore to successfully counterbalance the tendency toward a cascade phenomenon in the realm of public knowledge, the excesses made possible by an unrestrained scale-free network need to be counterbalanced through either one of two mechanisms: either a reduction in the number of connections afforded by the very few, or an increase in the density of the local network for individual entities. Either of these approaches may be characterized under the same heading: the fostering of diversity. The mechanism for attaining the reliability of connective knowledge is fundamentally the same as that of attaining reliability in other areas; the promotion of diversity, through the empowering of individual entities, and the reduction in the influence of well-connected entities, is essentially a way of creating extra sets of eyes within the network. This leads to the statement of the semantic condition (Figure 3):

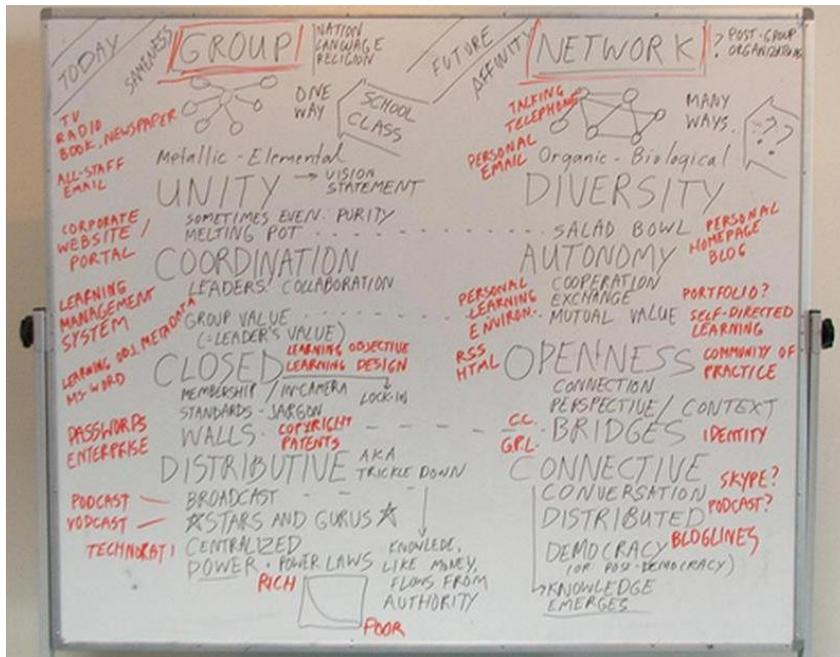


Figure 3. The distinction between groups and networks.

- First, *diversity*. Did the process involve the widest possible spectrum of points of view? Did people who interpret the matter one way, and from one set of background assumptions, interact with people who approach the matter from a different perspective?
- Second, and related, *autonomy*. Were the individual knowers contributing to the interaction of their own accord, according to their own knowledge, values and decisions, or were they acting at the behest of some external agency seeking to magnify a certain point of view through quantity rather than reason and reflection?
- Third, interactivity, or *connectedness*. Is the knowledge being produced the product of an interaction between the members, or is it a (mere) aggregation of the members' perspectives? A *different* type of knowledge is produced one way as opposed to the other. Just as the human mind does not determine what is seen in front of it by merely counting pixels, nor either does a process intended to create public knowledge.
- Fourth, and again related, *openness*. Is there a mechanism that allows a given perspective to be entered into the system, to be heard and interacted with by others?

## A NETWORK PEDAGOGY

The diagram in the previous section distinguishes between 'networks' and 'groups'. While it may be tempting to take this as a statement of some sort of ontology ('the world is divided into networks and groups, and these are their essential characteristics') it is better to think of the two categories as *frames* or *points of view* from with one may approach the creation of learning environments. After all, the same words may be used to describe the same entities at the same time, and this reflects not an error in categorization but rather the gestalt nature of the distinction. If the network theory applies to individual minds as well as to societies, then the network pedagogy I am proposing may be summarized as follows (and I *know* it's not original, or even substantial enough to be a theory properly So Called):

---

Downes Educational Theory
A good student learns by practice, practice and reflection.
A good teacher teaches by demonstration and modeling.
The essence of being a good teacher is to be the sort of person you want your students to become.
The most important learning outcome is a good and happy life.

---

While this may not appear to amount to much on the theoretical side, it – in combination with the four elements of the semantic condition – amounts to a robust pedagogy.

In essence, on this theory, to learn is to immerse oneself in the network. It is to expose oneself to *actual* instances of the discipline being performed, where the practitioners of that discipline are (hopefully with some awareness) *modeling* good practice in that discipline. The student then, through a process of interaction with the practitioners, will begin to *practice* by replicating what has been modeled, with a process of *reflection* (the computer geeks would say: back propagation) providing guidance and correction.

Learning, in other words, occurs in communities, where the practice of learning is the participation in the community. A learning activity is, in essence, a *conversation* undertaken between the learner and other members of the community. This conversation, in the web 2.0 era, consists not only of words but of images, video, multimedia and more. This conversation forms a rich tapestry of resources, dynamic and interconnected, created not only by experts but by all members of the community, including learners.

Probably the greatest misapplication of online community in online learning lies in the idea that a community is an adjunct to, or follows from, an online course. This is perhaps most clearly exemplified by the existence in itself of course discussions. It is common to see the discussion community created with the first class and disbanded with the last. The community owes its existence to the course, and ends when the course does. But the relation ought to be the other way around: that the course content (if any) ought to be subservient to the discussion, that the community is the primary unit of learning, and that the instruction and the learning resources are secondary, arising out of, and only because of, the community.

What needs to be understood is that learning environments are multi-disciplinary. That is, environments are not constructed in order to teach geometry or to teach philosophy. A learning environment is an emulation of some 'real world' application or discipline: managing a city, building a house, flying an airplane, setting a budget, solving a crime, for example. In the process of undertaking any of these activities, learning from a large number of disciplines is required.

These environments cut across disciplines. Students will not study algebra beginning with the first principles and progressing through the functions. They will learn the principles of algebra as needed, progressing more deeply into the subject as the need for new knowledge is provoked by the demands of the simulation. Learning opportunities - either in the form of interaction with others, in the form of online learning resources (formerly known as learning objects), or in the form of interaction with mentors or instructors - will be embedded in the learning environment, sometimes presenting themselves spontaneously, sometimes presenting themselves on request.

The idea of context-sensitive learning is not new. It is already supported to a large degree in existing software; Microsoft's help system, for example, would be an example of this were the help pages designed to facilitate learning and understanding. Cross (2006) is talking about a similar thing when he talks about informal learning. In a similar manner, learners interacting with each other through a learning environment will access 'help' not only with the software but also with the subject matter they are dealing with. Learning

will be available not in learning institutions but in any given environment in which they find themselves.

The Personal Learning Environment (PLE), which has attracted a lot of discussion in recent months, ought to be seen in this light. It is tempting to think of it as a content management device or as a file manager. But the heart of the concept of the PLE is that it is a tool that allows a learner (or anyone) to *engage* in a distributed environment consisting of a network of people, services and resources. It is not *just* Web 2.0, but it is certainly Web 2.0 in the sense that it is (in the broadest sense possible) a *read-write* application.

Attwell (2006) writes, “The promise of Personal Learning Environments could be to extend access to educational technology to everyone who wishes to organise their own learning. Furthermore the idea of the PLE purports to include and bring together all learning, including informal learning, workplace learning, learning from the home, learning driven by problem solving and learning motivated by personal interest as well as learning through engagement in formal educational programmes.” (§ 11).

The ‘pedagogy’ behind the PLE – if it could be still called that – is that it offers a portal to the world, through which learners can explore and create, according to their own interests and directions, interacting at all times with their friends and community. “New forms of learning are based on trying things and action, rather than on more abstract knowledge. ‘Learning becomes as much social as cognitive, as much concrete as abstract, and becomes intertwined with judgment and exploration.’” And – crucially – *teaching* becomes the same thing as well. As I wrote in 2002, “Educators play the same sort of role in society as journalists. They are aggregators, assimilators, analysts and advisors. They are middle links in an ecosystem, or as Hiler puts it, parasites on information produced by others. And they are being impacted by alternative forms of learning in much the same way, for much the same reasons.” (§ 10).

## **POSTSCRIPT: THE NON-CAUSAL THEORY OF KNOWLEDGE**

In recent years we have heard a great deal about evidence based educational policy. It is an appealing demand: the idea that educational policy and pedagogy ought to be informed by theory that is empirically supported. Such demands are typical of causal theories; following the methodology outlined theorists like Carl Hempel ([www.iep.utm.edu/h/hempel.htm](http://www.iep.utm.edu/h/hempel.htm)), they require an assessment of initial conditions, an intervention, and a measurement of observed difference, as predicted by a (causal) generalization.

In the earlier theory, there is a direct causal connection between states of affairs in the communicating entities; it is, therefore, a causal theory. But in the latter theory, there is no direct causal connection; it is what would be called (in the parlance of the new theory) an emergentist theory (that is, it is based on emergence, not causality). Calls for “evidence that show this claim is true” and “studies to substantiate this claim” are, like most Positivist and Positivist-inspired theories, reductive in nature; that is why, for

example, we expect to find something like a reductive entity, 'the message', 'the information', 'the learning', and the like. They are also aggregationist; the presumption, for example, is that knowledge is cumulative, that it can be assembled through a series of transactions, or in more advanced theories, 'constructed' following a series of cues and prompts.

But what happens, first of all, if the entities we are 'measuring' don't exist? Even if there are mental states, it may still be that our descriptions of them nonetheless commit some sort of category error. Saying that there are 'thoughts' and 'beliefs' that somehow reduce to physical instantiations of, well, *something* (a word, a brain state...) is a mistake. These concepts are relics of an age when we thought the mental came in neat little atomistic packages, just like the physical. They are an unfounded application of concepts like 'objects' and 'causation' to phenomena that defy such explanation; they are, in other words, relics of 'folk psychology'. Saying 'someone has a belief' is like saying that 'the Sun is rising' - it is literally untrue, and depends on a mistaken world view. But even more significantly, what happens if we cannot 'measure' the phenomena in question at all?

On the network theory knowledge and learning are emergent phenomena, and it is necessary to highlight a critical point: emergent phenomena are not causal phenomena. That is (say) the picture of Richard Nixon does not 'cause' you to think of the disgraced former president. They require a perceiver, someone to *recognize* the pattern being displayed in the medium. And this recognition depends on a relevant (but not strictly defined) similarity between one's own mental state and the pattern being perceived. That's *why* perception (and language, etc), unlike strict causation, is context-sensitive. And there is no means for a student to 'cause' (strictly speaking) recognition on the part of, say, an examiner, that he or she 'knows that Paris is the capital of France'. What is essential (and, I might add, ineliminable) is that the complex of this person's behaviours be recognized as displaying that knowledge. As Wittgenstein says, we recognize that a person believes the ice is safe by the manner in which he walks on the ice. And because this demonstration is non-causal, it depends on the mental state of the examiner, and worse, because (quite literally) we see what we want to see, the prior disposition of the examiner.

If this is the case, the very ideas of 'evidence' and 'proof' are turned on their heads. "Modeling the brain is not like a lot of science where you can go from one step to the next in a chain of reasoning, because you need to take into account so many levels of analysis... O'Reilly likens the process to weather modeling." (ScienceDaily, 2006, October 6, ¶ 12). This is a very important point, because it shows that traditional research methodology, and for that matter, traditional methods of testing and evaluation, as employed *widely* in the field of e-learning, will not be successful (are high school grades a predictor of college success? Are LSAT scores? Are college grades a predictor of life success?). This becomes even more relevant with the recent emphasis on 'evidence-based' methodology, such as the Campbell Collaboration ([www.campbellcollaboration.org/](http://www.campbellcollaboration.org/)) . This methodology, like much of the same type, recommends double-blind tests

measuring the impacted of individual variables in controlled environments. The PISA samples are an example of this process in action ([www.pisa.oecd.org](http://www.pisa.oecd.org)).

The problem with this methodology is that if the brain (and hence learning) operates as described by O'Reilly (and there is ample evidence that it does) then concepts such as 'learning' are best understood as functions applied to a distributed representation, and hence, will operate in environments of numerous mutually dependent variables (the value of one variable impacts the value of a second, which impacts the value of a third, which in turn impacts the value of the first, and so on).

As I argue in papers like *Public Policy, Research and Online Learning* (2003) and *Understanding PISA* (2005) the traditional methodology fails in such environments. Holding one variable constant, for example, *impacts* the variable you are trying to measure. This is because you are not merely screening the impact of the second variable, you are screening the impact of the first variable on itself (as transferred through the second variable). This means you are *incorrectly* measuring the first variable.

Environments with numerous mutually dependent variables are known collectively as chaotic systems. Virtually all networks are chaotic systems. Classic examples of chaotic systems are the weather system and the ecology. In both cases, it is not possible to determine the long-term impact of a single variable. In both cases, trivial differences in initial conditions can result in significant long-term differences (the butterfly effect).

This was a significant difference between computation and neural networks. In computation (and computational methodology, including traditional causal science) we look for specific and predictable results. Make intervention X and get result Y. Neural network (and social network) theory does not offer this. Make intervention X today and get result Y. Make intervention X tomorrow (even on the same subject) and get result Z.

This does not mean that a 'science' of learning is impossible. Rather, it means that the science will be more like meteorology than like (classical) physics. It will be a science based on modeling and simulation, pattern recognition and interpretation, projection and uncertainty. One would think at first blush that this is *nothing* like computer science. But as the article takes pains to explain, it *is* like computer science - so long as we are studying *networks* of computers, like social networks.

Learning theorists will no longer be able to study learning from the detached pose of the empirical scientist. The days of the controlled study involving 24 students ought to end. Theorists will have to, like students, immerse themselves in their field, to encounter and engage in a myriad of connections, to immerse themselves, as McLuhan would say, as though in a warm bath. But it's a new world in here, and the water's fine.

## REFERENCES

- Attwell, G. (2006). *Personal learning environments*. Retrieved November 23, 2008, from [http://www.knownet.com/writing/weblogs/Graham\\_Attwell/entries/6521819364](http://www.knownet.com/writing/weblogs/Graham_Attwell/entries/6521819364)
- Bach, M. (2003). Critical View on a Concept to Present Learning Material using Different Didactical Theories in a Learning Environment. In *Computers and Advanced Technology in Education*.
- Buchanan, M. (2002). *Nexus: Small worlds and the groundbreaking science of networks*. New York, NY: W. W. Norton & Company.
- Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind-brain*. MIT Press.
- Churchland, P. S., & Sejnowski, T. J. (1992). Computation in the age of neuroscience. In D. Hillis (Ed.), *The New Computation*. MIT Press,
- Cross, J. (2006). *Informal Learning: Rediscovering the Natural Pathways That Inspire Innovation and Performance*. Pfeiffer.
- Downes, S. (2002). *Aggregators, assimilators, analysts and advisors*. Retrieved November 23, 2008, from <http://www.downes.ca/cgi-bin/page.cgi?post=84>
- Downes, S. (2003). Public policy, research and online learning. *ACM Ubiquity Views*, 4(25). Retrieved November 23, 2008, from [http://www.acm.org/ubiquity/views/v4i25\\_downes.html](http://www.acm.org/ubiquity/views/v4i25_downes.html)
- Downes, S. (2004). *The Buntine oration: Learning networks*. Retrieved November 23, 2008, from <http://www.downes.ca/cgi-bin/page.cgi?post=20>
- Downes, S. (2005). *E-learning 2.0*. Retrieved November 23, 2008, from <http://www.elearnmag.org/subpage.cfm?section=articles&article=29-1>
- Downes, S. (2005). Understanding PISA. *Turkish Online Journal of Distance Education-TOJDE*, 6(2). Retrieved November 23, 2008, from <http://tojde.anadolu.edu.tr/tojde18/articles/article10.htm>
- Dretske, F. (1999). *Knowledge and the flow of information*. Center for the Study of Language and Inf.
- Dreyfus, H., & Dreyfus, S. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*, Oxford, U.K.: Blackwell.
- Fodor, J. (1983). *RePresentations: Philosophical essays on the foundations of cognitive science*. The MIT Press

- Fodor, J. (1989). *Psychosemantics: The problem of meaning in the philosophy of mind*. The MIT Press.
- Fodor, J. (2005). *The language of thought*. Harvard University Press
- Garson, J. (2007) *Connectionism*. Retrieved November 17, 2008, from <http://plato.stanford.edu/entries/connectionism/>
- Griffin, E. (2002). *A first look at communication theory* (5<sup>th</sup> edition). McGraw-Hill.
- Hanson, N. R. (1958). *Patterns of discovery: An inquiry into the conceptual foundations of science*. Cambridge University Press.
- Hebb, D. O. (2002). *The organization of behavior: A neuropsychological theory*. Lawrence Erlbaum.
- Hinton, G. E. (1986). Learning distributed representations of concepts. In *Proc. Ann. Conf. of the Cognitive Science Society*, volume 1.
- Huff, D. (1993). *How to lie with statistics*. New York, NY: W. W. Norton & Company.
- Jha, A. (2005). *Where belief is born*. Retrieved November 20, 2008, from <http://www.guardian.co.uk/science/2005/jun/30/psychology.neuroscience>
- Kirkpatrick, D. (2006). *Life in a connected world*. Retrieved November 20, 2008, from <http://money.cnn.com/2006/06/23/technology/brainstormintro.fortune/index.htm>
- Kosslyn, S. M. (2002). *Image and mind* (Reprint edition). Harvard University Press.
- Lakoff, G. (1990). *Women, fire, and dangerous things*. University Of Chicago Press.
- LeDoux, J. (2002). *Synaptic self: How our brains become who we are*. Viking Adult.
- Lewis, D. K. (2001) *Counterfactuals* (2nd edition). Wiley-Blackwell.
- Madden, M., & Fox, S. (2006). *Riding the waves of "Web 2.0": More than a buzzword, but still not easily defined*. Retrieved November 23, 2008, from [http://www.pewinternet.org/pdfs/PIP\\_Web\\_2.0.pdf](http://www.pewinternet.org/pdfs/PIP_Web_2.0.pdf)
- Malpas, J. (2005). *Donald Davidson*. Retrieved November 17, 2008, from <http://plato.stanford.edu/entries/davidson/>
- Marr, D. (1982). *Vision*. Times.

- Milligan, C. (2006). *What is a PLE? The future or just another buzz word?* Retrieved November 23, 2008, from [http://www.elearning.ac.uk/news\\_folder/ple%20event](http://www.elearning.ac.uk/news_folder/ple%20event)
- Minsky, M., & Papert, S. (1969). *Perceptrons: An introduction to computational geometry*. The MIT Press.
- Moore, M. G. (1973). Towards a theory of independent learning and teaching. *Journal of Higher Education*, 44(9), 661-679.
- Moore, M. G. (1989). Editorial: Three types of interaction. *The American Journal of Distance Education*, 4(2), 1-6.
- Moore, M. G. (1993). Three types of Interaction. In K. Harry, M. John, & D. Keegan (Eds.), *Distance education: New perspective*. London: Routledge.
- Nagel, T. (1974), What is it like to be a bat? *Philosophical Review*, 435-50. Retrieved November 3, 2008, from <http://www.clarku.edu/students/philosophyclub/docs/nagel.pdf>
- Newell, A., & Simon, H. A. (1963). GPS, a program that simulates human thought. In E. A. Feigenbaum and J. Feldman, *Computers and Thought* (pp. 279 – 293). McGraw-Hill. Retrieved November 18, 2008, from <http://www.math.grinnell.edu/~stone/events/scheme-workshop/gps.html>
- O'Hear, S. (2006). *e-learning 2.0: How Web technologies are shaping education*. Retrieved November 23, 2008, from [http://www.readwriteweb.com/archives/e-learning\\_20.php](http://www.readwriteweb.com/archives/e-learning_20.php)
- O'Reilly, R. C. (2006). Modeling integration and dissociation in brain and cognitive development. In Y. Munakata & M.H. Johnson (Eds) *Processes of Change in Brain and Cognitive Development: Attention and Performance XXI*. Oxford University Press.
- O'Reilly, T. (2005). *What is web 2.0: Design patterns and business models for the next generation of software*. Retrieved November 23, 2008, from <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- Polanyi, M. (1974). *Personal knowledge: Towards a post-critical philosophy*. University Of Chicago Press.
- Quine, W. V. O (1951). Two dogmas of empiricism. *The Philosophical Review*, 60, 20-43. Retrieved November 20, 2008, from <http://www.ditext.com/quine/quine.html>
- Quine, W. V. O (1964). *Word and object*. The MIT Press.

- Ravenscroft, I. (2004). *Folk psychology as a theory*. Retrieved November 3, 2008, from <http://plato.stanford.edu/entries/folkpsych-theory/>
- Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (1987). *Parallel distributed processing: Foundations*. The MIT Press.
- Scott, W., Oleg, L., Phil, B., Colin, M., Mark, J., & Paul, S. (2006). *Personal Learning Environments: Challenging the dominant design of educational systems*. Retrieved November 23, 2008, from <http://hdl.handle.net/1820/727>
- Searle, J. R. (1980) Minds, brains, and programs. *Behavioral and Brain Sciences* 3(3), 417-457. Retrieved November 3, 2008, from <http://www.bbsonline.org/Preprints/OldArchive/bbs.searle2.html>
- Siemens, G. (2004). *Connectivism: A learning theory for the digital age*. Retrieved November 23, 2008, from <http://www.elearnspace.org/Articles/connectivism.htm>
- Stalnaker, R. C. (1987). *Inquiry*. The MIT Press.
- Stephenson, K. (1998). What knowledge tears apart, networks make whole. *Internal Communication Focus*, 36. Retrieved November 23, 2008, from <http://www.netform.com/html/icf.pdf>
- Schramm, W. (1964). *Mass media and national development*. Stanford, CA: Stanford University Press.
- Schramm, W. (1997). *The beginnings of communication study in America: A personal memoir*. Thousand Oaks, CA: Sage.
- Technorati (2006). *State of the blogosphere*. Retrieved November 20, 2008, from <http://www.sifry.com/alerts/archives/000436.html>
- van Fraassen, B. C. (1980). *The scientific image*. Oxford University Press, USA.
- Watts, D. J. (2004). *Six degrees: The science of a connected age*. New York, NY: W. W. Norton & Company.
- Weinberger, D. (2002). *Small pieces loosely joined*. Cambridge, MA: Perseus Books.
- Wittgenstein, L. (1973). *Philosophical investigations* (3rd Edition). Prentice Hall.
- University of Colorado at Boulder (2006, October 6). Human Brain Region Functions Like Digital Computer. *ScienceDaily*. Retrieved November 23, 2008, from <http://www.sciencedaily.com/releases/2006/10/061005222628.htm>