EVOLUTION, CATEGORIZATION AND VALUES

Simon Winter

Lund University Cognitive Science
Kungshuset, Lundagård
S–222 22 LUND
Sweden

simon.winter@fil.lu.se

Abstract: The aim of this paper is to present an evolutionary framework for categorization. Evolution needs an evaluation mechanism to work, and it is argued that primary values that the organism needs for its survival – such as food, mates for reproduction, and shelter – can drive the evolution of categories. Sensory stimulation is needed to build up the cognitive apparatus, but cannot in itself provide the evaluation mechanism for evolution. Categorization constrained by values will be dependent on the availability of sensory information, and its power as predictive of values. As perception and categorization are tied to the actions of the organism, it is argued that the unit of perception should be seen as larger than the usual single-dimension stimulus, and evidence is reviewed to support this claim. Covarying stimuli will also provide a much greater predictive power than single-dimension stimuli alone.

1. INTRODUCTION AND OVERVIEW

The present paper provides an evolutionary framework for discussing some fundamental features of perception and cognition and tries to point out a number of current controversies in cognitive science that can be resolved by adopting this perspective. The main questions are: What can provide a firm ground for a theory of categorization, if we insert it into an evolutionary framework? Is it enough to build a theory of categorization on sensory input alone, and what place does "reality" have in such a theory?

An evolutionary perspective on categorization[1] provides natural constraints on what categorization can be like. First and foremost, evolutionary theories need some form of *evaluation mechanism* – categorization has to be *about* something for evolution to work, for example *about* finding food or avoiding danger. These evaluation mechanisms I have called values.[2] The organism needs substances with *food value* for its survival, needs to find a partner with *reproduction value* to produce offspring, must protect itself from predators with *destruction value*. These are *objective* constraints that all living systems have respected throughout their evolutionary history. One of the major themes in this paper is to present values as the driving force of categorization (section 2).

Thus, categorization is built upon cognitive and sensory processes, but is *about*, for example, finding food, which is not used for feeding cognition, but for feeding the body. The value systems constitute the *life processes*. Categorization is *about* keeping the creature alive. Eating is not primarily a cognitive process, with the aim of stimulating the senses in the digestive system, but consists of uniting substances possessing food value with our body to maintain the homeostasis. See figure 1.[3]

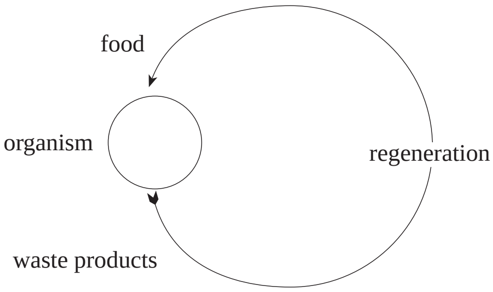

Figure 1. The value loop for substances with food value.

These physiological processes are (rightly) taken for granted in the literature of psychology and philosophy, but the distinction between cognitive processes and life processes cannot be overlooked in an evolutionary account, as the life processes constitute the

[1] Early works on evolutionary epistemology include Campbell (1974/1987) and Lorenz (1973/1977).

[2] Mainly in accordance with Gibson (1979). See further below. This term is close to von Uexküll's (1982) concept of "meaning."

[3] Cf. the "functional circle" of von Uexküll (1982).

Lund University Cognitive Studies – LUCS 66 ISSN 1101–8453. © 1998 by the author.

mechanism of evaluation of the evolutionary process. One of the main themes of this paper is to account for the connection between cognitive processes and life processes, or, as I call it, *sense domains* and *value domains*.

Consider, as an example, the two meanings of the word *hurt*. It refers both to the *sensation of pain*, and to the *physiological process of injury*.[4] Thus, there is a close connection in our minds between sense processes and value processes. And the only means we have of escaping injury is to escape the pain! How would we be able to protect ourselves if the sensory impulses did not correctly predict the values? In section 2.1, I examine the correspondence between sense domains and value domains.

There are some other nontrivial consequences of considering the closed loop of the life processes. Like any other cognitive feat, categorization must for example respect the time limits set up by the survival of the organism. Thus, categorization can be seen as a trade-off between the availability of information and its predictive power. See section 2.2.

Another consequence is that all categorization is seen as embodied and situated. The cognitive functions are seen in relation to the functions of the organism, in a context. The actions of the organism always take place in a multidimensional environment. In this environment, a multitude of information is available for use as sensory information, whether or not it is used by the organism. The bacteria in the Petri dish on my desk have for example access to the same potential of sensory information as I have. They could look at me as I look at them, but they don't.

This condition, that all organisms are surrounded by a vast potential of sensory information, is used to challenge the common view that categorization is built up from inferences in *single dimensions*. Rather, in section 3, I argue that the basic units of categorization are complexes of *covarying properties*. The unit of perception will thereby lie closer to the unit of interaction.

The phenomenon of categorization is ascribed in the literature to a vast range of organisms, from protozoa to humans. Informally, it may described as follows:

> given the motivational state of the organism, it has to find useful situations with food, mates and shelter, and must avoid danger, such as enemies, cliffs or excessive sunlight. In each situation, there is a choice to be made as to how to proceed, and this choice represents the categorization of the organism.[5]

There are of course several ways of performing categorization. Finding my way out of a house of horrors, I can proceed by reasoning to work out where I came in, following the sensory information in the trails in the dust or "blindly" reacting to the weak daylight coming through the shuttered windows. A dog could use the trails and the daylight. An amoeba only the daylight.

In most theories of human categorization, it is said to serve *reasoning*. (See Komatsu (1992) for a review.) In principle, there is nothing wrong in assuming that the main function of *human* categorization and concept formation is to serve reasoning. But if we want to attain a deeper understanding of fundamental human cognition, if we want to understand the connection between animal and human cognition, or build artificial systems with categorization capabilities, then categorization must be based on something that can support evolutionary constraints, and does not rely on the advanced abilities of humans, notably language.

According to some traditions, humans use linguistic faculties for all forms of cognitive processing – the so-called Language of Thought (Fodor/Pylyshyn 1988). I have chosen not to follow this tradition, but rather to find the foundations of categorization in nonlinguistic cognition. This will provide the common ground for human and nonhuman categorization that an evolutionary approach will need.

Although I devote this paper to aspects of categorization that are general to all organisms, there are a number of important features of *human* cognition that are necessary to remember as uniquely human when discussing categorization: peeling off the cognitive characteristics of language will uncover underlying similarities between animal and human cognition. In this paper, I will limit myself to laying a stable ground for categorization in the value systems that are common to all organisms.

2. VALUES, SENSES AND BRAINS

In this section I go deeper into the concept of values and show its relation with perception and cognition.[6] The method I use is akin to those of Jennings (1906) and von Uexküll (1909/1985; 1982). I construct a world with some basic properties, and this will provide the starting point for the discussion of some themes of perception and cognition.

---

[4]Likewise in other languages, e.g. *göra ont* in Swedish.

[5]It is always the *situation* that must be categorized as beneficial or detrimental, not a singular object. For example, if the animal finds good food in a dangerous place, it must have a means of judging advantages against disadvantages. Not all animals are good at this, when manipulated by "smarter" animals. A mouse-trap is a familiar example.

[6]Readers acquainted with constructivism will recognize most of the terms and arguments, but will perhaps be surprised at the stress on *reality* as an important part of the theory. Constructivist literature includes von Glasersfeld (1976; 1977; 1984; 1995), Stewart (1996), Sjölander (1995; 1997/in press), Watzlawick (1984).

The world I assume has to be stable to some extent. At least, the aspects of the world that are captured by cognitive generalizations *have to be* stable, otherwise there would be no generalizations to capture! The organisms inhabiting the world – descendants of the Vehicles of Braitenberg (1984), and the Berry creatures of Gulz (1991) and Balkenius (1995) – are theoretical creatures, but exhibiting many behaviors common to living organisms.[7]

All creatures have to comply with the restrictions imposed by the closed loop of survival: intake of food, reproduction, homeostasis of temperature and other conditions of life, such as oxygen content, air pressure etc. For the lowest organisms, the sensory requirements are at a minimum: for bacteria in a heap of dung, no senses or sensory information are needed to find the food – it is always there. In higher organisms a radical shift has taken place due to symmetry breaking (Stewart/Golubitsky 1992). We eat at regular *intervals*, rather than continously, and as long as the intervals are not greater than the loop admits, we have the freedom to leave the food and do other things. But then we will need sensory organs to find the food again.

The world we move around in is a complex world. The theoretical analogy that I want to use is that of a *cave* that is not known to us in advance. We move around in the cave, and as long as we keep clear of the walls, nothing prevents our exploration. This world is *objective* in an ontological sense. It poses *constraints* on us that we have to comply with (Stewart 1996). I have illustrated this in figure 2.[8]

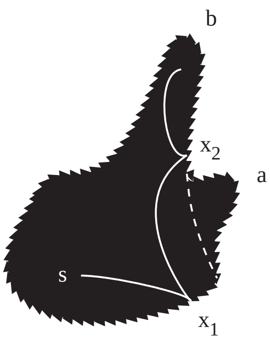

Figure 2. The organism starts at *s*, "hits reality" at x_1 and x_2.

The black area in the figure represents the world, in terms of objective possibilities: wherever the creature tries to go, inside the area, there will be nothing (in these two dimensions) that will stop it. The hinder that the "end of the world" represents has nothing to do with sensory stimulation. It is purely a matter of physical constraints.

The movement does not have to be restricted to *spatial* movement. Also other aspects of our experience can be represented dimensionally (Gärdenfors 1996). When we look for something to eat, we explore an area of potential food that extends over several dimensions. Some food is beneficial to us and will continue to be part of our world. Some food is dangerous and will stop our exploration.

The values that substances have will vary depending on species. This does not alter the physical characteristics of the substances, however, only their relation to the organism, and what aspects the organisms are likely to include in their mental representation. Thus, I argue that the substances with value do not in themselves need to be mentally represented. Rather, mental representation is concerned with strategies for finding and avoiding these substances.

As an illustration of the difference between sense domains and value domains, please put your hand on your desk. Push downwards. Further. When you sense the pain, ignore it and push harder. Did you get through? No. Again, we take for granted the association of the sense dimensions and what I have called the value dimensions – what prevents your hand from passing through the table is not the pain, it is the physical characteristics of the table, which correspond to the edges of the black world in the figures.

The organism in figure 2 starts its trajectory at *s*. When it comes to x_1 it "hits reality." The knowledge gained from this experience depends on the sensory apparatus of the creature. The one in the figure seems to have noticed the wall, but couldn't predict it. It changes direction, hits the "wall" again at x_2, and we leave it for the moment.

Irrespective of the objective layout of the world, the creature will, if it is complex enough, try to establish its own subjective map showing which areas are allowed. A system of inductive heuristics will allow the creature to extend its mental map based on its experience, as pictured in figure 3. The white area represents the parts of the space that the animal will treat *as if* it were safe.

Inside the white areas the animal is less likely to check carefully. It can switch from the slower attended, context-dependent mode of processing required when the informational predictability is low, to automated processing which is more rapid and error-free and can be performed in parallel with other tasks (Givón 1989: ch. 7).

[7]As von Uexküll (1909/1985) observed, when discussing the inner worlds of lower organisms, we as researchers always adopt the perspective of the organism in question. I will therefore allow myself to use the pronoun "we" even when I talk of bacteria and other lower organisms.

[8]For the sake of familiarity, I will assume that the dimensions are spatial during the discussion.

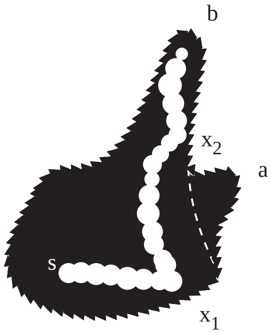

Figure 3. The mental map of a timid organism. The white area represents the subjective safe zone.

We do not so far have to assume anything about the perception of the organism, it could create a representation from dead reckoning (Gallistel 1990), based on the trajectory. But it could also use, for example, landmarks or smell gradients (Balkenius 1995; 1996).

What says, then, that the creature should delimit its expected subjective harmful zone to a small area around the experienced trajectory? Figure 4 shows the mental map of an "epistemically bold" creature.

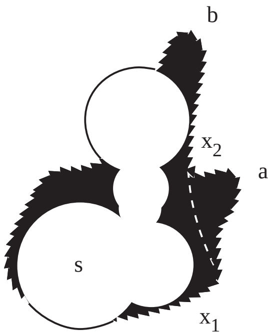

Figure 4. The white space represents the mental map of a bold organism.

Here, the creature still avoids the two places x_1 and x_2 of the "reality encounters." (This is the creature's response to the stability requirement.[9]) In this case, the creature has drawn too optimistic inferences that have extended *beyond* the limits of the objectively "possible." However, as long as it does not try these possibilities, it will never get the negative feedback from reality, and will not have to revise its inferences.[10] Furthermore, the bold creature has the advantage of inductively knowing a far greater portion of the allowed space than has the timid creature, and this might counterbalance the negative effects of over-generalization.[11] As I have tried to illustrate by the number of intersecting circles, the creature also gains *simplicity* in the representation.[12]

2.1. "Direct perception" and the function of the senses

I would like to compare my approach above with the theory of affordances of Gibson (1979). A succinct characterization is given by Neisser (1987):

> *Affordances*, as J. J. Gibson (1979) defined them, are relations of possibility between animals and their environments. A particular environment has a given affordance if and only if it makes a given kind of action possible, whether that action is actually executed or not. The claim that a given affordance exists is an objective claim, always either true or false: I may or may not be able to walk on that surface, for example.

This is the first part of Gibson's affordance theory, and the least discussed in a general framework – the *objective* character of affordances. Of course, and Gibson also notes this, the affordances are species-specific. Thin ice supports a mouse but not a cow. The affordances correspond to the black areas – "reality" – in the figures above.

The second part of Gibson's argument is that affordances are *directly perceived*, and this I would like to discuss further.[13] I have tried above to evoke a picture where the sensory dimensions are separated from the value dimensions, and I would like to follow that line of reasoning for a while before returning to Gibson again.

If we assume direct perception of affordances, this would yield a perfect match between inner and outer environment, as in figure 5. Gibson studied mainly visual perception, and for vision it might be easier to obtain a "true image" of the environment, although some parts of the cave are obscured from certain vantage points.

---

[9] The limits of reality do not generally move. "Agents" are exceptions (Givón 1989; Premack 1996).

[10] "Reality is what makes your expectations fail." (Per Johansson, pc)

[11] Depending on the negative feedback in the dimension in question.

[12] One possibility is to see the bold creature as generalizing on a coarser scale (Balkenius 1996).

[13] This is the most common interpretation of Gibson's thoughts, and the part that provides the foundations of the theory of "visibility" of Norman (1988) – we act *directly* on cues in the environment.

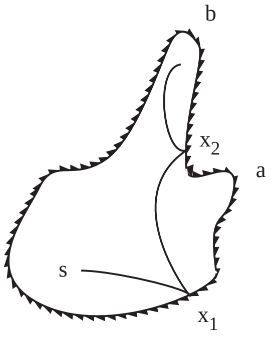

Figure 5. The mental map of an informed organism. Perfect correspondence between inner and outer worlds.

As I said before, the space that the creature investigates does not have to be spatial. Take a common example: the creature has to find food that is edible by exploring two food dimensions. The black space in the figures now represents the objectively safe zone, and the organism has to find sensory cues to get to know the space. Now, the negative feedback is harder: a "reality hit," i.e. eating something outside the borders, causes illness or death. The problem is that there is no a priori correspondence between the sensory dimensions and the value dimensions. Or is there?

Sensory organs represent an inductive heuristic that if we base our actions *in the real environment* upon the categories that we can find *in the sensory input,* then we survive. The "only" problem is that there are very many worlds compatible with our sensory input. Evolution, however, has helped us solve this problem of the adaptation of the senses to our environment:[14]

> The single individual is still a prisoner in his constructed world, but the system as such will slowly, over millions of generations, improve its correspondence between what goes on inside the brain and outside it (Sjölander 1995).

Gibson's theory is directly appealing in that it stresses the way we tie perception to *action* without reflection. Perception is not seen as passive intake of information into a storage unit, but as a direct guide to action. This thought also makes it possible to link human perception to earlier evolutionary stages, for example down to bacteria, which base their locomotion on sucrose gradients (Stewart 1996).

As humans we envision the problem as one of picking mushrooms in the forest. A combination of visual and olfactory cues guide us, and neither the visual nor the olfactory information provides information enough to form a true map of the space of edible mushrooms.

For lower organisms, there is the direct link from perception to action, which in some sense makes the border between the sense domains and the value domains disappear. Thus, Gibson's theory predicts a state where it is possible to *take the sense information for the value information* – i.e. there is a correspondence between the *salience* in a sensory dimension and the values that we need for survival. We can react directly to the sensory stimuli. We don't spit out "bitter" food. We spit out "dangerous" food. This is a heuristic that saves much cognitive effort, and as long as the species survives it is a viable heuristic.[15] For higher organisms, however, the case is more complicated, as we have cognitive structures mediating between perception and action (see section 2.3).

Evolution has helped us evolve sense organs that allow us to make distinctions corresponding to useful divisions of our environment. However, there are several situations where the correspondence does not hold.

> Fruits and berries, for example, have more food value when they are ripe, and this is specified by the color of the surface (Gibson 1979:131).

In the quotation from Gibson, we see again how the association of the sense domain and the value domain is taken for granted. But just because we cannot take the association for granted, the correspondence between sensory properties and values is not completely arbitrary. In the case of fruits and berries there *are* for example chemical reactions in the value domain that change color when fruits grow ripe and have more food value. Some changes in skin color in fruits and berries, on the other hand, can be seen as a property in the sense domain that has co-evolved with animal *use* of the property. Certain plants depend on birds for disseminating their seeds, and thus the sensory stimulus that the birds use as a cue will be reinforced by evolution.

The preceding discussion touches the old problem of our direct access to Kant's "Ding an sich," and provides an evolutionary answer to the question of whether we can get to know the world directly: Insofar as the sensory organs really have adapted to the outer reality, what we get from them is knowledge about reality. Unfortunately, there does not seem to be any fundamental way of knowing *to what degree this adaptation has taken place*. However, the only conclusion granted by the theory of evolution is of course that the sensory organs have adapted to a degree where they have let the organisms survive in the specific environment where they have lived. Whether this means that we are exploring the borders of the cave, or stay somewhere in the middle, remains an open question.

[14] This idea originates in Lorenz (1973/1977:6–7).

[15] As a heuristic it is quite fixed, and an *evolutionary* heuristic rather than an individual one!

2.2. Availability vs. predictive power

The constitution of our sense organs makes some properties more *available* than others. The eye is sensitive only to a certain frequency range of light, the ear to another range of sound. Some substances lack smell for our species. This is of course not because there is an absence of molecules of the substances to be picked up by our olfactory system, but because we have not adapted to this particular substance during evolution.

On the other hand, it is not certain that the most available properties will lead us to the right place in the value space: their *predictive power* is not necessarily very great.

To determine the utility of categorization, I propose to see categorization as a trade-off between availability and predictive power. In fact, this is a *consequence* of discussing categorization in an evolutionary framework: it is necessary to adopt a pragmatic point of view, where categorization is coupled with action, and this real-world action will be the mechanism of evaluation in evolution. Due to the limitations imposed by the maintenance of life processes, all categorization in nature will be bounded.

So, on the one hand we must investigate what information is readily available to us, through our sensory organs, and on the other we must see what kind of information we need for different purposes.

Availability will thus in practice put a limit on categorization in many situations. Many edible mushrooms are left uneaten in the forest, not because it is not *possible* to distinguish them from poisonous ones, but because the information needed is not *readily available*. The categorization procedure will change depending on the degree of *risk* involved.

2.3. Senses vs. brains – the emergence of concepts

We might use olfactory sense information to distinguish an edible mushroom from a poisonous one, and we do this because there are generalizable situations where it works. The introduction of a "nicely" smelling but poisonous mushroom into such a situation will force us to look for another distinguishing property, either a finer olfactory distinction, or a distinction in another domain, such as the visual.[16] All such refinements will need additional representational feats: we will need some primitive "concept."

[16]Downplaying the most salient attributes of objects is equivalent to postulating the existence of *nonobvious* properties as important for categorization. A commonly used example is the distinction between bird and bat, where overall similarity is outweighed by genetic relation. See the brief discussion of nonobvious properties in Gelman/Coley (1991).

The concept is an intermediary layer between sensory input and value domains (see figure 6).[17] Concepts in this sense will comprise everything from a temporary downplaying of salience, due to attention, to human cognition.

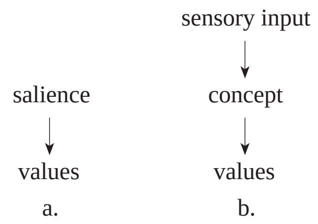

Figure 6. In some cases, we can act directly upon salient sensory input. In others, salience is downplayed by concepts.

One proposed model to consider here is the *subsumption architecture* by Brooks.[18] Cognitive representations ("concepts") are seen as layered modules that can suppress lower layers, for example sensory input. There is also an element of competition between different modules, giving the cognitive system a certain amount of flexibility.

2.4. Extending the world of values – conditioning

To take but one example of the dichotomy between categorization with and without concepts, let us briefly compare the approach of this paper with *conditioning*. Conditioning is a term borrowed from behaviorism, traditionally accused of disregarding the role of internal representations. However, as we will see, it is also one of the few theories to account for the immediate association of the sensory domain and the value domain.[19]

In classical conditioning, there is a fundamental distinction between two kinds of stimulus. The first kind, for historical reasons called *unconditioned stimulus* (US), is directly connected to the value domains, and thus has *meaning*[20] for the animal. Unconditioned stimuli include food, which makes the animal salivate (Pavlov 1927), and various stimuli causing pain, such as electric shocks.

The other stimuli in the environment do not have this connection to the value domains, but are never-

[17]In this very general discussion, I have chosen the term "concept" for the intermediary layer, although as will soon be clear, some of the concepts are *very* rudimentary. Other terms that could be used are "theory," "schema," or "representation."

[18]Brooks (1991), reviewed in Balkenius (1995).

[19]See Balkenius (1995:ch. 5) for an overview of conditioning, and Rescorla (1988) for some relatively late developments within the behaviorist tradition. Following a suggestion by Christian Balkenius (pc), it is possible to see conditioning as a *performance test* rather than a foundational property of cognition. This view makes conditioning easy to reconcile with representational theories.

[20]Especially in the sense of von Uexküll (1982).

theless salient: flashing lights, ringing bells and the like were used in the experiments.

It turned out that if a meaningless stimulus is presented in connection with a meaningful one, the animal forms some kind of association – the formerly meaningless bell becomes a predictor of the meaningful food. The bell functions as a *conditioned stimulus* (CS). After a sequence of trials, the connection has grown so strong that the CS alone produces the response. Observe that the foundation of this form of learning always rests on the association of stimuli with the value domain. This is a way for the animal to *create meaning*.

From one perspective, the association of an unconditioned stimulus with a response must be seen as direct, as in figure 6 a, and the conditioned stimulus producing a response as mediated, as in figure 6 b. However, as for example the behaviorists maintained, there is perhaps no need to postulate cognitive representations at all to describe these processes. This question will remain open.[21]

2.5. Evolutionary essentialism

In the literature, the *essence* of categories is often discussed as properties of an instance that must be present for the instance to be an instance of that category (Gelman/Coley 1991). An albino tiger is still a tiger, as long as it possesses tiger DNA, for example. The essence in essentialist theories provides an *evaluation* of the category, but is not postulated against some criterion *outside* the system of categorization.

In my framework, where the *raison d'être* for the creation of categories is that they orient us towards the value domain, the essence will be given by the association with the value domain. The only thing that we require of a category in its simplest form is that it points the way to primary values.[22] Thus, an evolutionary essentialism will not require our *knowing* what the essences *are* – they are chiseled out by the evolutionary pressure on the categories.

Depending on whether a category represents something that an animal wants to get or to avoid, we will have to distinguish two cases. A category that represents a *positive* primary value, i.e. something that the individual must have for its survival, will always have an "essence," otherwise the animals with this kind of categories die. If, for example, we base our survival on a cereal that does not give us essential amino acids, we will not survive.[23]

Categories representing *negative* primary values will not be subject to the same constraints: we can fear something that is not dangerous without ever getting negative feedback on this categorization. We can continue avoiding the dark cave in the forest, even though the dragon died many years ago – we will never know if we don't see for ourselves. Thus, for these categories we do not have to postulate that they are grounded in an essence, only that the categorizations are subject to constraints for reasons of cognitive economy.

3. THE UNIT OF INTERACTION AND THE UNIT OF PERCEPTION

I have pictured a scene that is common to all living beings, where the basic primitives are the survival loop, and the substances with food, danger and reproduction value. Values are not cognitive, and when we benefit from them, it is not basically from sensory stimulation, although even the simplest organisms have evolved senses as a *guide* to values.

The next question in this paper is *in what form* we conceive of these substances – as single stimuli clustering to form objects, as holistic objects being decomposed into dimensions, as whole environments or as combinations of the three. I will not reach a definitive conclusion on these matters,[24] I want rather to put the searchlight on some evidence that has existed for a while but that has not received the attention it deserves. Although the evidence is inconclusive it comes from a range of cognitive disciplines, and that might help as a corroboration.[25]

3.1. A complex unit of perception

As I argued above, the essential interaction of the organism with its environment is in the *value* domains. Due to the physical properties of organism

---

[21]As a general epistemological standpoint: in scientific developments there often arises the need for a *dialectical contrast* between positions. The content a concept has will be dependent on what it is *contrasted against*. If we discuss human concepts, we may say that they clearly mediate between perception and values, and then an association like conditioning will in this simple model be direct. See Andersson (1994) for this kind of polemic concept formation.

[22]In higher-level cognition there are of course also second-order categories that are not coupled to a primary value domain. A simple example is money. It is not useful in itself, but we "reward ourselves" when we get it.

[23]Medin/Ortony (1989) advocates a view almost contrary to mine: psychological essentialism, described as "not the view that *things* have essences, but rather the view that people's *representations* of things might reflect such a belief (erroneous as it may be)." Also Gelman/Medin (1993:163): "Essences are typically not known, almost always unobservable, and may not exist. So, the essence itself cannot usually serve as the basis of how people categorize or identify items." However, when studying human categorization that is connected to language and intentionality, this position is more tenable.

[24]The debate on category coherence is long and continuing. See for example Keil (1989), Millikan (to appear), Quine (1969), Gelman/Coley (1991), Medin (1983).

[25]Early developments of this question are discussed in Campbell (1966:89). Cf. also the tradition of Gestalt psychology (Köhler 1947), and the excellent critique in Leyton (1992) in terms of reduction of Gestalts to causal histories with the aid of symmetry principles.

and environment, this interaction is complex – it is not possible to imagine the interaction as an interface only letting through one dimension at a time. Furthermore, there is often much *potential sensory information* in the organism's environment, that the organism might use as a guide to the value domains.

Given this complexity in the environment, I will argue that the best level of description of perception is on a more complex level than in most classical psychology and philosophy. This is in accordance with the characterization by Campbell (1966:82):[26]

> Both psychology and philosophy are emerging from an epoch in which the *quest for punctiform certainty* seemed the optimal approach to knowledge. To both Pavlov and Watson, single retinal cell activations and single muscle activations seemed more certainly reidentifiable and specifiable than perceptions of objects or adaptive acts.

From an evolutionary perspective it is arguable that the unit of perception would lie closer to the unit of interaction. I will show evidence of representations at least on an intermediate level between the "objects" of interaction and the more primitive analytical level common in psychology. One such example is the concept of affordances in Gibson's (1979) ecological theory of perception:

> [...] what we perceive when we look at objects are their affordances, not their qualities. We can discriminate the dimensions of difference if required to do so in an experiment, but what the object affords us is what we normally pay attention to. [...] If this is true for the adult, what about the young child? There is much evidence to show that the infant does not begin by first discriminating the qualities of objects and then learning the combinations of qualities that specify them. Phenomenal objects are *not* built up of qualities; it is the other way around. The affordance of an object is what the infant begins by noticing. The meaning is observed before the substance and surface, the color and form, are seen as such. *An affordance is an invariant combination of variables, and one might guess that it is easier to perceive such an invariant unit than it is to perceive all the variables separately* (Gibson 1979:134–5, my emphasis).

”Invariant combination of variables” should be read as variables that *covary* with each other, but not with other properties. I will therefore examine the concept of covariation over the next subsections.

[26] Linguistic semantics has suffered the same decomposition into minimal features.

3.2. The covariation heuristic

Even our most basic practices involve complex sensory input. As I have discussed in section 2.2, categorization is dependent on the predictive power of the available sensory information. However, if the organism can detect *covariation* in the sensory input, the rudiments of concepts can be formed. The various combinations of stimuli that are obtained when covariation is taken into account have a much greater predictive power than the stimuli in isolation. This heuristic can be characterized as one of inductive inference:

> Inference from clustering of categorial properties:
>
> (a) "Individual members of a natural category do not share only a single criterial property. Rather, they most often share many properties, which are thus the *definitional core* of their categorial membership."
>
> (b) "Therefore, if known members of a group exhibit properties A, B, C etc., and if a sample sub-group also exhibits property Z (to a statistically-significant degree), then it is *highly likely* that the rest – untested – members also exhibit property Z." (Givón 1989:276)

However, it seems to be clear that not all animals utilize all possible covariation information. As reported in Sjölander (1995), snakes, for example cannot use cross-modal covariation. Snakes use different modalities for completing different subtasks, such as catching the prey and swallowing it. This should not be considered as odd from an evolutionary point of view. In the same way as not all potential sensory information is used by animals, there is a great potential of covariation information that *could* be used. However, it is likely that different sensory stimuli will provide a basis for covariation for different organisms. Covariation detection will be an *economical* solution for categorization, but only up to a certain breakpoint, which will be decided by the evolution of the organism.

Cross-modal covariation can be used by human infants to integrate different senses. Piaget (1968/1970:90) relates the case of an infant of 3 months at the developmental stage when the visual system is gaining autonomy. There is a functioning "réaction circulaire" consisting of hand movements mainly towards the mouth. The child uses hand movements as a source of covariation in connection with vision: the hand can be moved with relative autonomy, and the first step for the visual system is to follow the movements of the hand. The opposite is not yet true, so the infant cannot look at an object, and then reach out and grasp it, a faculty that it will soon achieve. Thus, there is a possibility to integrate sensory information by means of covariation with a sense that already has meaning in that it can be used to obtain values. Vision, working at a distance, is not so easy to integrate with value loops, and thus needs the initial support of the haptic sense.

3.3. Basic level objects

In psychology we find another account that parallels Gibson's approach.[27] Rosch (1978:31) presents the *basic level* of categorization:

> A working assumption of the research on basic objects is that (1) in the perceived world, information-rich bundles of perceptual and functional attributes occur that form natural discontinuities, and that (2) basic cuts in categorization are made at these discontinuities.

I believe that in the present evolutionary framework it is safe to extend her framework from the "perceived" world to the "real" world. We could not have an evolution of a conceptual system without a counterpart in the value domain.[28] (Given this firm evolutionary conviction, it would also be possible to try to reconcile the psychologistic position in Rosch's account with the realist discussion about *natural kinds*.[29])

3.4. Perception of complex properties

What then is the evidence to corroborate that we do in fact detect covarying properties directly, rather than by composition from more primitive sense domains? From semantics, for example, we are used to thinking about concepts as decomposable into primitive features that are then processed bottom–up to form complex concepts by production rules.[30] My argument in this section is that this is not the most fruitful level of explanation, since it has no connection to the value domains. Rather, we will look for representations that correspond more closely to the level of interaction of the animal. Tanaka (1993) shows that in the anterior infero-temporal cortex (TE), we find *single-cell activation of complex* features:

> The critical features were more complex than orientation, size, color, and simple texture, which are known to be extracted by cells in V1. Some of the features were shapes that were moderately complex, whereas others were combinations of such shapes with color or texture. The individual critical features were not complex enough to specify a particular object seen in nature through activation of a single cell. Activation of a few to several tens of cells with different critical features seems necessary to specify a particular natural object (Tanaka 1993:686).

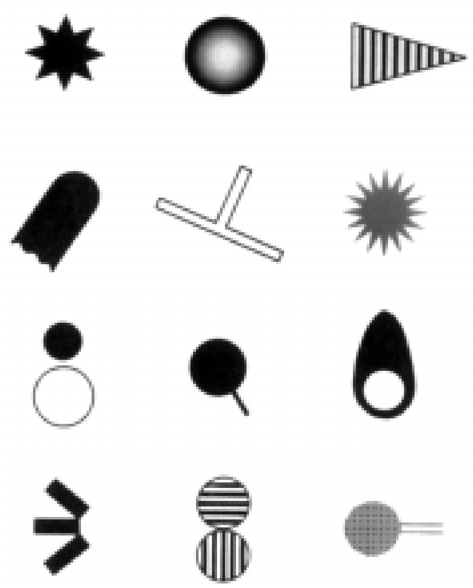

Figure 7. Twelve examples of critical features for the activation of single cells in area TE. From Tanaka (1993). Some features are also colored in the original.

Figure 7 shows some of the critical features. The importance of Tanaka's results is that they challenge our view of features as being describable as simple concepts in Euclidean geometry that generate more complex representations by production rules.

Given the stimuli that the visual system must distinguish, *various* intermediate levels could be imagined. It is very hard to predict the stimulus set that a combination of such a kind of "moderately complex" features would be able to distinguish, and even more impossible to find the *best* intermediate set for a given set of stimuli. But what could be learned from this is (a) that there exists a level of representation that does not benefit from being decomposed into primitives, and (b) that this level is not generated by geometrical primitives, but by evolutionary pressure on real-world categorizations.[31,32]

[27]There are also other theoretical developments in various traditions. Shanks et al. (1996), following a suggestion by Rescorla, talk about *configural* stimuli: co-occurring stimuli should be treated as unique and not as the simple sum of their parts.

[28]What we of course cannot say anything about is whether there are *other values* that could be utilized with another conceptual system, or whether there are *other combinations of sensory domains* that could provide a better foundation for categorization.

[29]Richard Boyd (1991) provides similar evidence, based on ideas indicating that natural kinds should be viewed as *homeostatic property clusters*: this is the same idea about covariation, supplemented with a claim regarding category coherence by homeostasis. There is a striking resemblance between this account of homestasis and the description of life processes in terms of "autopoiesis" in Maturana/Varela (1987/1992).

[30]E.g. Lyons (1977:ch. 9).

[31]See also Lettvin et al. (1959) for a discussion of the receptors in the frog's eye, and that "the language in which

The theory that I have developed above proposes that categorization and perception are closely tied to our interaction with elements in the environment, and in particular those elements that carry values with a direct impact on ourselves. As humans we share the need for values with lower organisms, and we are as dependent as they are upon food, mates and shelter. But we have also evolved higher representational skills that are not closely tied to our basic interaction with the world.

4. CONCLUSIONS

In the paper, I have provided an evolutionary framework to open up a discussion about some fundamentals of categorization and cognition. As a summary, I would give only a few points of reference to the issues discussed.

• The organism needs substances with primary value for its survival. These substances are not *stimuli* in that their important function is to stimulate the senses of the organism. Rather the organism *merges* with them. Values represent the *essence* of the categories.

• Senses are our means of making contact with the substances that have value for us. Together with conceptual representations they let us diverge from the essential substances for a time and find them again more efficiently afterwards, but "we must respect our meals," and be careful not to break the loop of life processes.

• Evolution has shaped our senses so that for *some* categories we can react directly to our sense impressions *as if* they represented the essence/value. This explains what is sometimes called direct perception.

• There are both beneficial and harmful values. There is an imbalance in how we can react toward them. We can continue to avoid things that are not dangerous, but we have discontinued eating things that made us die.

• There is a moment in time when the categorization takes place, and it is signaled by the *behavior* of the individual. The categorization has to be done with the available sensory information that does not necessarily carry the optimal predictive power. Thus, there is a trade-off between availability and predictive power.

• In many theories of categorization, categories are built up from singular dimensions. In the current approach, the fact that we always *act* in complex environments is exploited as an advantage. The unit of perception will reflect the unit of interaction. In the organism–value loop there is always a multitude of information available that can provide a basis for sensorical inferences.

• Both the theory of values and the theory of covariation originate in the assumption that we interact with real-world objects rather than with singular sensory stimuli.

• A theory of values is necessary to maintain the evolutionary continuity between animal and human categorization, since symbolic (or memetic) evolution is subject to partly different evolutionary constraints.

ACKNOWLEDGMENTS

This project was financed by Lund University. I want to thank Christian Balkenius, Per Johansson, Annika Wallin and Peter Gärdenfors for the inspiration for many of the ideas in this paper, and Tom Andersson, Sverre Sjölander, David de Léon, Lars Kopp, Linus Broström and the Lund University Cognitive Science group for their help and enthusiasm.

REFERENCES

Andersson, T. (1994) Conceptual Polemics – Dialectic Studies of Concept Formation. Ph.D. thesis, Cognitive Science Dept. Lund University, Lund.

Balkenius, C. (1995) Natural Intelligence in Artificial Creatures. Ph.D. thesis, Cognitive Science Dept. Lund University, Lund.

Balkenius, C. (1996) Generalization in Instrumental Learning. In P. Maes, et al. (eds), From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior, Cape Cod, Massachusetts, The MIT Press.

Balkenius, C./Kopp, L. (1996) The XT-1 Vision Architecture. In P. Linde/G. Sparr (eds), Symposium on Image Analysis, Lund 1996.

Billman, D. (1983) Procedures for Learning Syntactic Categories – A Model and Test with Artificial Grammars. Ph.D. dissertation, University of Michigan.

Billman, D./Heit, E. (1988) Observational Learning From Internal Feedback – A Simulation of an Adaptive Learning Method. Cognitive Science 12 (4): 587–625.

Boyd, R. (1991) Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds. Philosophical Studies 61: 127–148.

Braitenberg, V. (1984) Vehicles – Experiments in Synthetic Psychology. The MIT Press, Cambridge, MA.

Brooks, R.A. (1991) How to Build Complete Creatures Rather than Isolated Cognitive Simulators. In: K. VanLehn (ed), Architectures for Intelligence. Lawrence Erlbaum Ass., Hillsdale, NJ, pp. 225–239.

Campbell, D.T. (1966) Pattern Matching as an Essential in Distal Knowing. In: K.R. Hammond (ed), The Psychology of Egon Brunswik. Holt, Rinehart and Winston, New York, pp. 81–106.

---

they are best described is the language of *complex abstractions from the visual image*. We have been tempted, for example, to call the convexity detectors 'bug perceivers.'" (p. 1951, my emphasis)

[32] There is some literature explicitly dealing with covariation detection, e.g. Kareev (1995), Billman (1983), Billman/Heit (1988), Nisbett/Ross (1980), but most of the stimuli used in their experiments require language, and linguistically based representations are subject to other evolutionary laws than the ones discussed in this article (Dennett 1995/1996:ch. 12).

Campbell, D.T. (1974/1987) Evolutionary Epistemology. In: G. Radnitzky/W.W. Bartley (eds), Evolutionary Epistemology, Rationality, and the Sociology of Knowledge. Open Court, La Salle, IL, pp. 47–89.

Dennett, D.C. (1995/1996) Darwin's Dangerous Idea - Evolution and the Meanings of Life. Penguin Books, London.

Fodor, J./Pylyshyn, Z. (1988) Connectionism and cognitive architecture: a critical analysis. In: S. Pinker/J. Mehler (eds), Connections and symbols. MIT Press, Cambridge, MA.

Gallistel, C.R. (1990) The Organization of Learning. The MIT Press, Cambridge, MA.

Gelman, S.A./Coley, J.D. (1991) Language and Categorization: The Acquisition of Natural Kind Terms. In: S.A. Gelman/J.P. Byrnes (eds), Perspectives on Language and Thought – Interrelations in Development. Cambridge U P, Cambridge, pp. 146–196.

Gelman, S.A./Medin, D.L. (1993) What's So Essential About Essentialism? A Different Perspective on the Interaction of Perception, Language, and Conceptual Knowledge. Cognitive Development 8: 157–167.

Gibson, J.J. (1979) The Ecological Approach to Visual Perception. Houghton Mifflin, Boston.

Givón, T. (1989) Mind, Code and Context – Essays in Pragmatics. Lawrence Erlbaum Ass., Hillsdale, NJ.

Gulz, A. (1991) The Planning of Action as a Cognitive and Biological Phenomenon. Ph.D. thesis, Cognitive Science Dept. Lund University, Lund.

Gärdenfors, P. (1996) Conceptual Spaces as a Framework for Cognitive Semantics. In: A. Clark/J. Ezquerro/J.M. Larrazabal (eds), Philosophy and Cognitive Science – Categories, Consciousness, and Reasoning. Kluwer, Dordrecht, pp. 159–180.

Jennings, H.S. (1906) Behavior of the lower organisms. Columbia University Press, New York.

Kareev, Y. (1995) Positive Bias in the Perception of Covariation. Psychological Review 102 (3): 490–502.

Keil, F.C. (1989) Concepts, Kinds, and Cognitive Development. The MIT Press, Cambridge, MA.

Komatsu, L.K. (1992) Recent Views of Conceptual Structure. Psychological Bulletin 112 (3): 500–526.

Köhler, W. (1947) Gestalt Psychology. Liveright, New York.

Lettvin, J.Y./Maturana, H.R./McCulloch, W.S./Pitts, W.H. (1959) What the Frog's Eye Tells the Frog's Brain. Proceedings of the Institute of Radio Engineering of New York 47: 1940-1951.

Leyton, M. (1992) Symmetry, Causality, Mind. The MIT Press, Cambridge, MA.

Lorenz, K. (1973/1977) Behind the mirror – A search for a Natural History of Human Knowledge. Methuen, London.

Lyons, J. (1977) Semantics – Volume I. Cambridge U.P., Cambridge.

Maturana, H.R./Varela, F.J. (1987/1992) The Tree of Knowledge – The Biological Roots of Human Understanding. Shambhala, Boston.

Medin, D./Ortony, A. (1989) Psychological Essentialism. In: S. Vosniadou/A. Ortony (eds), Similarity and Analogical Reasoning. Cambridge U. P., Cambridge, pp. 179–195.

Medin, D.L. (1983) Structural Principles in Categorization. In: T.J. Tighe/B.E. Shepp (eds), Perception, Cognition, and Development – Interactional Analyses. Lawrence Erlbaum Ass., Hillsdale, NJ, pp. 203–230.

Millikan, R.G. (to appear) A Common Structure for Concepts of Individuals, Stuffs, and Real Kinds – More Mama, More Milk and More Mouse. Behavioral and Brain Sciences.

Neisser, U. (1987) From direct perception to conceptual structure. In: U. Neisser (ed), Concepts and conceptual development: Ecological and intellectual factors in categorization. Cambridge University Press, Cambridge, pp. 11–24.

Nisbett, R./Ross, L. (1980) Human Inference – Strategies and Shortcomings of Social Judgment. Prentice-Hall, Englewood Cliffs, NJ.

Norman, D.A. (1988) The Design of Everyday Things. Doubleday, New York.

Pavlov, I.P. (1927) Conditioned Reflexes. Oxford, London.

Piaget, J. (1968/1970) La naissance de l'intellligence chez l'enfant. Delachaux et Niestlé, Neuchâtel.

Premack, D. (1996) Cause/induced motion: intention/spontaneous motion. In: J.-P. Changeux/J. Chavaillon (eds), Origins of the Human Brain. Clarendon Press, Oxford, pp. 286–308.

Quine, W.V. (1969) Natural Kinds. In: W.V. Quine (ed), Ontological Relativity and Other Essays. Columbia University Press, New York, pp. 114–138.

Rescorla, R. (1988) Pavlovian Conditioning – It's Not what you Think. American Psychologist 43: 151–160.

Rosch, E. (1978) Principles of Categorization. In: E. Rosch/B.B. Lloyd (eds), Cognition and Categorization. Lawrence Erlbaum Ass, Hillsdale, NJ, pp. 27–48.

Shanks, D.R./Lopez, F.J./Darby, R.J./Dickinson, A. (1996) Distinguishing Associative and Probabilistic Contrast Theories of Human Contingency Judgment. In: D.R. Shanks/K.J. Holyoak/D.L. Medin (eds), Causal learning (The Psychology of Learning and Motivation, vol. 34). Academic Press, San Diego, pp. 265–311.

Sjölander, S. (1995) Some cognitive breakthroughs in the evolution of cognition and consciousness, and their impact on the biology of language. Evolution and Cognition 1 (1): 3–11.

Sjölander, S. (1997/in press) Fulguration, Dezentrierung und inneres Probehandeln - Evolutionäre Durchbrüche in der Evolution der Sprache und Kommunikation, im Sinne von Lorenz, Piaget und Freud. In: G. Guttmann/F.M. Wuketits (eds), Freud–Piaget–Lorenz. Wiener Studien zur Wissenschaftstheorie.

Stewart, I./Golubitsky, M. (1992) Fearful Symmetry – Is God a Geometer? Blackwell's, Oxford.

Stewart, J. (1996) Cognition = Life: Implications for higher-level cognition. Behavioural Processes 35: 311–326.

Tanaka, K. (1993) Neuronal Mechanisms of Object Recognition. Science 262 (29 October 1993): 685–588.

von Glasersfeld, E. (1976) The Development of Language as Purposive Behavior. In: S. Harnad, et al. (eds), Origins and evolution of Language and Speech. Annals of the New York Academy of Sciences. The New York Academy of Sciences, New York, pp. 212–219.

von Glasersfeld, E. (1977) Linguistic Communication – Theory and Definition. In: D.M. Rumbaugh (ed), Lanugage Learning by a Chimpanzee: The Lana Project. Academic Press, New York, pp. 55–71.

von Glasersfeld, E. (1984) An Introduction to Radical Constructivism. In: P. Watzlawick (ed), The Invented Reality – How Do We Know What We Believe We Know? Contributions to Constructivism. W. W. Norton, New York, pp. 17–40.

von Glasersfeld, E. (1995) Radical Constructivism: A Way of Knowing and Learning. The Falmer Press, London.

von Uexküll, J. (1909/1985) Environment [Umwelt] and Inner World of Animals. In: C.J. Mellor/D. Gove (eds), Benchmark. University of Tennessee, Knoxville.

von Uexküll, J. (1982) The Theory of Meaning. Semiotica 42 (1): 25–82.

Watzlawick, P. (ed). (1984) The Invented Reality – How Do We Know What We Believe We Know? Contributions to Constructivism. W. W. Norton, New York.