Aestheticism in the Theory of Custom

by

EKKEHART SCHLICHT

Journal des Économistes et des Études Humaines

2000, volume 10, number 1
pages 33– 51

AESTHETICISM IN THE THEORY OF CUSTOM*

Ekkehart Schlicht°

> First we may observe, that the supposition, that the future resembles the past, is not founded on arguments of any kind, but is derived entirely from habit., by which we are determined to expect for the future the same train of objects, to which we have been accustomed.
>
> David Hume (1740, 134)

1. Introduction

Customs, habits, and routines provide the bedrock for many economic and social formations yet our understanding of the processes that underlie the growth and decay of customs is very limited. The theory of social evolution has hardly commenced to evolve.

The 'clarity' view of custom proposed in my recent book *On Custom in the Economy* posits the desire of individuals to detect patterns in their social environment and to act in a patterned fashion. They have a 'rule preference', and this gives rise to the formation of customs and social evolution. In this essay, I offer some supplementary arguments which support this position from the perspective of learning theory and evolutionary psychology.

The first issue to be dealt with relates to learning (Sections 2 to 8). What processes should we envisage for the way in which the rules of custom are learned by individuals in a society? Obviously the rules of custom and social interaction must be learned. It is usually taken for granted that this learning proceeds in an adaptive way, and it is assumed that people find out one way or another what is best for them and adjust their behavior accordingly. Social experimentation goes on without respite. Competitive forces select more successful behavior and enforce

* Discussions with Eva Jablonka have been important in shaping the ideas presented here. I thank Peter Weise and an anonymous referee for penetrating comments.

° Professor of Economics, Department of Economics, University of Munich, Germany, e-mail: schlicht@lrz.uni-muenchen.de.

Volume 10, numéro 1, Mars 2000, pp 33-51.

it on the individuals. In the end, a social structure emerges from this process of incessant mutual re-adjustment. This picture, drawn notably by theorists dealing with social evolution, is, however, ambiguous.[1] It leaves the question open as to whether the rules of custom grow out of experience, or whether people experiment with alternative rules and select the best from the set. Both theoretical alternatives have been pursued in the literature in a cursory fashion, and without expanding on detail. I shall refer to the first view as 'rule inductivism' and to the second as 'rule structuralism'. Neither view can give an adequate account of rule formation and the processes underlying the assimilation of customary behavioral patterns (Section 2 to 5).

A close examination of rule learning reveals that learning processes are intimately tied up with evaluations of an aesthetic kind, relating to formal features like symmetry, analogy, or good continuity. This observation leads to a third alternative, rule-aestheticism, which takes the middle ground between inductivism and structuralism. It offers a more satisfactory account of the learning processes that channel social evolution and is described in Sections 6 to 8.

The issue of rule learning will be discussed in a very simple setting. I shall concentrate almost exclusively on conventions, which are rules that solve co-ordination problems and where it is best for each individual to follow the convention if others do the same. Keeping on the right hand side of the road is an example. Conventions, in contrast to many other prescriptions of custom, do not pose enforcement problems. This permits concentration on fundamental aspects of learning processes.

The processes of learning are of particular importance for the social sciences because the way people learn influences their behavior and thereby moulds the social regularities which are to be learned. In this sense, learning processes are of more fundamental significance in the social sciences than, say, in physics, where a physicist may neglect the fact that the way he thinks is part of physical reality and that the aesthetic judgements involved in generating his theories play any role in it. We may allow that we would approach problems in physics differently if we were endowed with another type of aesthetic sense, but we could still be confident that the theories thus developed would describe the same physical reality, and would amount, in this sense, to much the same as the theories we currently entertain. With regard to social structure, however, we must expect that another type of aesthetic sense would have made us settle for quite different property or family structures. We would live in a different social world.

2. Inductivism

Let me consider learning first, and how people learn the rules of custom. One view of rule formation posits that rules are formed inductively from experience. Adam Smith envisaged the formation of the rules of moral conduct in this manner:

[1] Boyd/Richerson-1985.

> *"Our continual observations upon the conduct of others, insensibly lead us to form to ourselves certain general rules concerning what is fit and proper either to be done or to be avoided. It is thus that the general rules of morality are formed. They are ultimately founded upon our experience of what, in particular circumstances, our moral faculties, our natural sense of merit and propriety, approve, or disapprove of. We do not originally approve or condemn particular actions; because, upon examination, they appear to be agreeable or inconsistent with a certain general rule. The general rule, on the contrary, is formed, by finding out from experience, that all actions of a certain kind, or circumstanced in a certain manner, are approved or disapproved of."*[2]

A similar position may be found in modern game theory where it is maintained that notions of fairness reflect and encode successful strategic behavior as learnt and adopted in a process of trial and error, with successful behavior maintained, and unsuccessful behavior avoided. Successful behavior is summarized in terms of rules. People follow these rules not necessarily because they are aware of their usefulness, but for emotional or moral reasons, yet these emotional or moral motives are only proximate causes of behavior. The motives themselves have been formed because they have generated successful behavior.[3]

According to rule inductivism, customs incorporate the inductively derived and emotionally encoded recipes for success. Analysis must penetrate the surface phenomena of moral preferences and judgements and zero in on the ultimate instrumental causes of customs. Any explanation of custom, it is maintained, must start from there.

3. Structuralism[4]

I would like to contrast the inductivist view sketched in the last section with another extreme view, labeled 'rule structuralism'. This view can be described as follows. There is a set of possible rules like 'you must not lie,' 'you must not steal,' or 'you must drive on the right-hand side of the road'. These rules are pre-fabricated ideas in a Platonic (or Kantian) 'rule-heaven,' given *a priori*. Humans select from this set by adopting certain rules, and rejecting others. The survival of certain rules, or certain rule-systems, can be analyzed again in competitive terms. The 'better' rules, or 'better' rule-systems, survive and supersede the others. Here 'better' means competitive dominance, or a faster spread of active rules within the population. The concept is akin to biological fitness.

2 Smith-1759, p. 159.

3 Binmore/Samuelson-1994, pp. 46-7.

4 The concept of structuralism is introduced and used here in a very simple way, *i.e.* by assuming that there is a set of pre-fabricated rules (structures) without any finer distinction among them. This is done in order to draw attention to the importance of making distinctions within the set of structures, as will be elaborated in section 6 below. Current structuralist positions in linguistics or the social sciences are, however, more refined and do take some of these aspect into account.

In the following I shall, however, not discuss issues of propagation and evolution of rule systems, but emphasize the crucial difference between rule inductivism and rule structuralism: The former takes rules as generated by competitive forces, the latter takes competition as taking place between rules. These rules must, therefore, precede competition.

The position of rule structuralism is shared by many theorists who think about the choice of alternative rule-systems, related to constitutional economics, or 'Wirtschaftsordnungen.'[5] Some positions in modern institutional economics may, in this sense, be classed as structuralist. They characterize institutions as 'humanly devised constraints that shape human interaction,' and interpret them as 'rules' which have been selected, or have emerged from competition.[6] Rule utilitarians can be counted as structuralists, too. They insist that general rules, rather than specific actions, are to be selected according to the results they bring about. This excludes the option of piecewise optimization.

4. A Critique of Inductivism[7]

Rule inductivism holds that rules are formed by induction from experience. This is, however, a position that is difficult to maintain because rules cannot emerge from rule-free or otherwise unaided induction. Obviously, induction requires ideas about how inductive knowledge is generated. Processes of induction can not be conceived as free-floating. They need an anchor.

Consider the following simple inference problem. An individual sets out to learn how to behave at various traffic crossings. He finds out that it is best to take action *a* under circumstances *A,* action *b* under circumstances *B,*
under circumstances *C*. Take the simple case of right of way in traffic, and take circumstances *A, B* and *C* as referring to particular crossings. The following characterizations of actions *a, b,* and *c* may be conceived:

a give right of way to cars coming from the right at crossing *A*
b give right of way to cars coming from the left at crossing *B*
c give right of way to boats crossing from leeward at crossing *C*

The individual notes also that crossing *A* is one particular crossing in Munich, crossing *B* is a crossing in London, and crossing *C* is a certain crossing of waterways on a inlet of the Baltic, near Kiel. Note that formation of the underlying notions 'right', 'left', 'windward,' or 'leeward' relies on induction in the sense that these characterizations and classifications must have been learned and that all individuals have adopted the same classifications. Furthermore, the individual must have learnt that the distinction windward-leeward is irrelevant on the road, and the distinction right-left is irrelevant on the waterway, and that a myriad of other possible

[5] Buchanan/Brennan-1985, Buchanan-1994.
[6] North-1990, p. 3.
[7] For the following, see also Goodman-1983, pp. 59-83 and Schlicht-1998, pp. 87-105.

distinctions are irrelevant at all crossings. Usually, many other characteristics, like north-south, or broad-narrow, may serve to co-ordinate action equally well. If each individual had tried to co-ordinate by using another characteristic, the learning of co-ordination would have been impossible. If one person tries to find a right of way rule, based on the windward-leeward distinction, while his partner concentrates on the right-left distinction, and a third individual tries to co-ordinate by concentrating on yet another aspect, it will be practically impossible for the group to co-ordinate successfully. Furthermore rule learning may involve generalizations such as 'in England, always give right of way to vehicles coming from the left, but on the continent, give right of way to vehicles coming from the right,' and 'stop to traffic from the leeward side on waterways'. Such generalizations rest on notions like 'England', 'the continent' or 'waterways'. These, again, must be shared by the individuals concerned and must precede any inductive learning of general rules.

All learning rests thus on classifications and distinctions. In order to learn co-ordination inductively, the relevant classifications and distinctions used by the individuals concerned must have been coordinated beforehand. The individuals must have settled spontaneously for matching characteristics as coordinating devices, or must have, at least, settled for a few possibilities, such as right-left or windward-leeward distinctions, in order to render learning possible. This set of alternatives can not be determined inductively, because the vast number of theoretical possibilities would frustrate any attempt to single out one particular distinction as being the relevant one. A lifetime would not suffice to gather the necessary information.[8]

Co-ordination can only emerge from mutually matching inductions drawn by the individuals concerned, and from a correlated cognitive structuring of experience. Ultimately, the possibility to learn and co-ordinate in social interaction rests on cognitive dispositions which are shared by the majority of individuals concerned. This fact is theoretically of great significance, but has been neglected in many theoretical inquiries. It is all too natural, like breathing, and does not stir up any attention. Yet, unlike breathing, the way we make inductions shapes social interaction. It cannot be ignored.

The argument that all learning is rule-bound (which differs, however, from the position developed in this paper) has been of great importance in linguistics regarding the acquisition of language and the learning of grammar rules. It seems to be well established that young children learn language guided by an innate knowledge of the possible forms of natural language. Without language universals, learning would not be possible, as each new achievement is generalized in many ways beyond whatever has been experienced before. Language acquisition must be understood in structuralist, rather than behaviorist, terms. This seems to be a widely accepted view. The debate today concentrates on the question whether the universals underlying language acquisition reflect language-specific knowledge or

[8] In linguistics, an analogous argument is used to defend the postulate of a generative grammar. Similar problems arise in biology with attempts to explain the evolution of behavior, but are often ignored. (See for example Maynard-Smith-1978).

general regularities of cognition. The idea of structuralist learning itself remains unchallenged.[9]

5. A Critique of Structuralism

The learning of the rules of custom is in many ways very similar to the learning of the rules of grammar. It involves many generalizations. If, at a certain crossing, we have learnt to give way to traffic coming from the right, we will spontaneously and subliminally form the rule 'right before left' and try this in other situations – at other crossings, on the sidewalks, etc. This type of generalization – that one learning event triggers off an entire cluster of other behaviors – is akin to the way we learn languages, and is essential for effective learning.

We have seen that inductivism cannot account for effective rule learning, However, structuralism cannot adequately elucidate rule-learning either. The simple dichotomy between rules and non-rules is not sufficient to cope with learning unless the possibilities are extremely and unrealistically limited.

To illustrate, consider the following four passing rules, which build on the right-left categorization, but add a temporal dimension:

R1 always keep right

R2 always keep left

R3 keep right on odd days (Mondays, Wednesdays, Fridays, and Sundays) and keep left on the other days of the week

R4 keep right on even days (Tuesdays, Thursdays, and Saturdays) and keep left on the other days of the week

When trying to learn a passing rule, we start with simple hypotheses like *R1* or *R2*. If these do not work, we may try refinements *R3* or *R4*. We would not, however, start with rule *R3* as a first approximation and then refine it by restricting it to odd days and invoke rule *R4* for the rest of the week, thereby effectively reproducing rule *R1* as a combination of rules *R3* and *R4*. This would appear unnatural and unwieldy. Yet theoretically it would be a matter of indifference whether we took *R1* and *R2* as our primitive rules, and conceived *R3* and *R4* as refinements, or whether we started from *R3* and *R4* and took *R1* and *R2* as refinements. If the only possible distinction refers to the one between rules and non-rules, all rules are equivalent. There would be no hierarchy in the rule heaven.

[9] See Anderson-1980, p. 352. The position of aestheticism developed in this paper could be applied to linguistics, too, and would then challenge certain aspects of linguistic structuralism. See also footnote 11 below.

Without a hierarchy, however, rule learning would not be realistically possible. The following example may clarify this thought. We consider the issue of driving on the right versus driving on the left and allow for rules that use distinctions of weekdays, just as rules *R1* to *R4* above do. Many other possible rules exist that build on the right-left distinction and on the classification of weekdays. All in all, 2^7=128 different rules can be stated. Each such rule may be described by a sequence of letters indicating the appropriate behavior on the corresponding day of the week. Thus we write *R1=(r, r, r, r, r, r, r)* or *R4=(l, r, l, r, l, r, l)*. If we allow for all possible rules, there would be no way to learn by induction from the past. Each observation on one day would cut future possibilities by half, because it would fix the choice of right or left for that particular day, but would not carry any implication for the remaining days of the week. Whatever had been observed at the beginning of the week will, thus, not help to make predictions for the remaining days.

Rule formation and rule learning will actually proceed differently. Assume that behavior of others has been observed on the first five days of the week. On each of the first five days, we have observed driving on the right. Our pre-conceived ideas of simplicity, clarity and straightforwardness would suggest to us to expect *(r, r)* for Saturday and Sunday, and we would confidently assume that the others were guided by similar expectations. This would enable smooth co-ordination on Saturday and Sunday, emerging from a generalization from past experience on the preceding weekdays. We would, so to speak, prefer the rule *R1=(r, r, r, r, r, r, r)* over the rule *R5=(r, r, r, r, r, l, l)* when making inductions. Without ideas of simplicity, clarity and continuity, however, no such grading of rules would be possible. The rule *R5* 'drive on the right save on Saturdays and Sundays' is a possible rule, just as rule *R1* 'drive on the right all the time'. Similar observations apply to all possible combinations of driving on the right and driving on the left.

We are, thus, able to learn from past experience, because we prefer certain inductions to others, and we happily assume that such a rule preference is a good guide for predictions about the future.[10]

Note that the induction problem has been discussed above in a very simple setting, assuming that other facts had been learnt before. For instance, it has been assumed that a 'week' is the relevant time period to consider, and that other than right-left categories do not matter. In a more realistic setting, the problem of determining which rule is best becomes practically insoluble.

It would be thus of no great help to restrict the set of rules to a subset, as rule structuralism would suggest. If we knew *a priori* that you should keep at the same side of the road on Saturdays and Sundays, this would no doubt restrict the set of possibilities by half. We would not need to learn about Sundays. However,

[10] See Schlicht-1998, Ch. 8. The argument relates closely to Goodman's contention (Goodman-1983). He points out that unaided induction is impossible, and proposes the view that the categorizations given in language serve this purpose. In a similar vein, and starting from the same problem, Goyal and Janssen propose that learning of conventions presupposes some other conventions (Goyal/Janssen-1996). The argument presented here would trace the emergence of these categorizations and conventions to clarity judgements of an aesthetic kind, ultimately prompted by our psychological make-up.

in order to render rule learning possible, the set of alternatives would have to be narrowed down drastically. Such a trimming of the rule heaven would be entirely unjustified on *a priori* grounds.[11] Every sequence of right and left could serve as a possible rule for coordinating passing on the road. It is only that some rules are considered better than others, in a purely aesthetic, non-instrumental sense. This induces us to try them out first. The position of structuralism ignores this fact and postulates, erroneously, that a clear-cut distinction can be drawn between rules and non-rules.

6. Aestheticism

Rules cannot be derived from unaided induction. Thus, inductivism, as conceived above, is an untenable position. Similarly, structuralism, as conceived above, cannot account for rule learning and rule formation, as it rests on an untenable categorical distinction between rules and non-rules. The examples given above suggest, however, an intermediate position that avoids both extremes and, at the same time, can account very naturally for rule formation and rule learning. This is the position of rule aestheticism, which will be described presently.

The basic observation here is that rules can be graded not only with respect to their instrumental usefulness, but also with respect to their clarity, straightforwardness, and ease of perception and reproduction. Some rules are better than others, in this sense. For the purpose of learning, induction, and transmission, individuals prefer more attractive to less attractive rules. They have a rule preference. This renders it possible to learn from the past.

Rule preference is of an essentially aesthetic nature. Symmetry, simplicity, straightforwardness, analogy, and other formal features contribute to distinguish a 'good' rule from a 'bad' one. The clarity of a rule is, however, not a number that can simply be attached to it, or springs from a calculation of the 'clarity values' of its components. Just like beauty 'is not in any of the parts or members of a pillar, but results from the whole,' the beauty or attractiveness of a rule depends on its overall pattern, and how well it fits in with other rules in the prevailing set of customs.[12]

This would show up empirically if we tried to measure the clarity of a rule by noting what types of rules people prefer and try out first. The rule of walking on the right on the sidewalk will appear more attractive on the continent than in Great Britain, because it would harmonize with the rules prevailing on the continent, not in Britain. Likewise, the rule to drive on the right on weekdays and on the left on Sundays would appear better than the rule to drive on the left on every other tenth

[11] The argument could also be advanced against structural linguistics: There is no clear-cut distinction to be drawn between correct and incorrect sentences. Some sentences are clear, some are murky, some verge on being wrong, and some are definitely wrong from a grammatical point of view.

[12] Hume-1777, p. 292. Let me note that the clarity preference is not to be equated to a preference for simplicity, see Schlicht-1998, p. 136.

day and on the right otherwise. This depends entirely on the prevailing convention of having a seven-day week, rather than a ten-day-week.

Aesthetic judgements, if shared by the individuals concerned, render it possible to solve the pervasive induction problem. Rule detection becomes possible because the 'better' rules are tried first, and modifications may only be introduced later if necessary.

This procedure is well illustrated in econometrics, where we start by assuming linear relations first. If we allowed for polynomials of arbitrary degree from the outset, there would be infinitely many which would fit our data perfectly, but there would be no way to decide which of these polynomials to choose. By assuming simple relationships first, and introducing modifications when needed, we can obtain our results.[13]

7. Deeper Levels

Both the inductivist and the structuralist approach can be refined and shifted to more fundamental aspects of learning. This gives rise to *sophisticated inductivism* on the one hand, and *sophisticated structuralism* on the other.

Sophisticated inductivism. It may be argued that the rules are formed by inductive processes on a higher level. What appears 'simple' or 'clear' *to us* is not simple or clear in any objective sense but is perceived as thus because it is advantageous to form this, and no other, notion of simplicity and clarity. Evolution has taught us to form such notions in the most expedient way. This argument points to a theoretical possibility but seems to me to be of limited bearing, at least in the context of the social sciences. We can safely assume that the fundamental processes of learning and behavior, which characterize humans, and are shared by many animals, are invariant in historical time. We can take mental structure as given.

On a pragmatic level, psychologists have addressed the issue of induction *vs.* pre-determined structure in concept formation. The inductivist position was that a concept – say, of a bird – refers to an average specimen which we most frequently

[13] This is the well-known identification problem in econometrics. If arbitrary functional forms are permitted for a regression equation, there will be infinitely possibilities to obtain a perfect fit for the past, with arbitrarily many associated predictions for the future. The problem is solved in econometrics by trying 'simple' functional forms (like straight lines, quadratic or logarithmic functions) first.

It is to be noted here that this is not simply a matter of the number of parameters involved, although the problem is usually discussed in this way. It is true that a linear equation $y=a+bx$ involves only the two parameters a and b, but this holds true for

$$y = a + b \sum_{i=2}^{100} \frac{x^i}{\log(i)}$$

as well, and infinitely many other two-parameter functions are conceivable.

In particular, for each set of observations $(x_t, y_t)_{t=1,2,\ldots T}$ and any prediction $(x_t, y_t)_{t=T+1,T+2,\ldots T+z}$ there will exist infinitely many polynomials $y = P(x)$ which will yield a prefect fit. Estimating the equation $y=a+b.P(x)$ will give the estimates $a=0$ and $b=1$, but this kind of perfect regression will tell us nothing about predictions because we can obtain all predictions we like in this way.

encounter, and we form the concept of a bird that fits best in most cases. The other alternative was that abstract features such as symmetry and clarity rather than frequent exposure or other practical concerns govern concept formation. It turns out that such abstract features and, in particular, the context in which observations occur, are very important for concept formation.[14] This is also evident in our everyday experience with the decimal system. The most prominent numbers here are 1, 10, 100, etc., but these are not the numbers we use most frequently. In the sexagesimal system which we use with timepieces, the numbers 60, 120, 180 and 240 are prominent. Such clarity judgements are driven by the number system in the first place, rather than by frequent exposure. Sometimes, frequent exposure is the result of, rather than the cause for, clarity features. We have, for example, television films which fit into 60 minute time-slots, and videotapes which are gauged to this rhythm.

Sophisticated structuralism. The observation that mental structure must be taken as fixed and given in historical time may suggest, again, a structural view of a more refined kind. Sophisticated structuralism forms its beginning from the idea that rule learning is rule-bound itself. Thus, it may be urged, there must be rules for learning rules. The 'deep' rules are genetically determined. They enable us to learn and to make inductions. This kind of structuralism could be developed in full analogy with linguistic structuralism. In linguistic structuralism, it is maintained that children are genetically equipped with a 'generative' grammar which enables them to learn any language which happens to be spoken by their caretakers in an extremely efficient manner. The generative grammar is, thus, a set of rules for making rules. If applied in full analogy to the social sciences, this kind of structuralism would maintain that humans are equipped with a 'generative social structure' which produces, in interaction with prevailing circumstances and historical conditions, any social structure we may observe.[15]

Sophisticated structuralism need not be conceived in such a modular manner, however. Just as cognitive dispositions enabling language acquisition may not be language specific, the cognitive dispositions enabling the learning of the rules of social interaction may be of a general nature, rather than specific to social interaction. We need neither postulate a separate language module nor assume a separate social module in our cognitive organization, as both language acquisition and social learning phenomena may stem from a general ability to learn rules.

8. Structuralism and Aestheticism

Structuralism, in its sophisticated non-modular version, distinguishes between a generative structure, genetically given, and the realized social structure

[14] See Anderson-1990, pp. 137-145 and Schlicht-1998, pp. 75-86 for further discussion.

[15] This would be one reading of Aristotele's 'hexis' or Pirker's and Rauchenschwandtner's 'sense of community', see Pirker/Rauchenschwandtner-1998, pp. 410-11.

or a rule system which we actually observe. A set of generative rules constitutes the 'deep structure'. It generates, in interaction with prevailing social and historical conditions, the particular rule systems we observe in different societies.

In contrast, aestheticism places great emphasis on the necessity of *grading* possible rules according to clarity and straightforwardness. It has been urged that a non-instrumental, or aesthetic, preference for clear rules must be presupposed. This rule preference induces people to try the clear rules first. This makes induction and rule learning possible. However, the grading of rules according to clarity has been described without referring to the different layers of rules such as deep generative and superficial actual rules. In this, the proposed view of rule learning deviates from structuralism. The position seems preferable for purposes of social analysis, as a distinction between different layers of rules is neither necessary nor simplifying. Furthermore, it is not obvious that a categorical distinction can usefully be made between generative and actual rules, as it seems that any rule, once adopted, may serve to generate other rules.

This is partially a semantic issue. Consider the case of rules to keep to one side on the footpath and on the street. Let *(r, l)* denote the case that you keep to the right on the footpath and to the left on the street. Assume a society with footpaths, but no streets, and where the rule was established to keep to the right on the footpath. With the introduction of carriages and carts, the necessity arose for streets and a rule for their use. The alternatives were, to select either *(r, r)* or *(r, l)* as a rule system. Rule preference would suggest the first alternative. Hence the previously established rule 'Keep to the right on the footpath' entails the derived rule 'keep to the right on the street'. In this sense, the first rule helped to generate the second. More generally, any rule can serve as a generative rule in so far as clarity judgements depend on context, and any rule can serve as an element in the context for the establishment of another rule. In this sense, a distinction between generative and superficial rules seems unwarranted.

We may phrase the same reasoning in terms of generative rules, however. The prescription: 'keep to the same side on the footpath and on the street' may be considered a generative rule in this case. As a matter of semantics, we may, in this vein, conceive that any principle which establishes a preference for a certain rule over another one, is a generative rule.

But semantic choices are rarely innocuous, as they ease certain types of arguments and impair others. In this sense, the semantic choice of distinguishing between generative and superficial structures seems unfortunate. This becomes evident in the cases where simplicity judgements and analogies are important. The rule 'go for simple rules' presupposes simplicity judgements, which could, in principle, be stated by some rules that describe the processes generating such judgements. These could then be taken as generative rules. In a similar vein, the rule 'treat similar cases analogously' can be traced to some generative rules which describe the way we form similarity judgements and analogies. Such a parlance in terms of generative structures seems unnecessarily cumbersome, however. It may be preferable to point directly to the types of judgement on which rule formation builds. These processes have been described here as 'aesthetic,' in the sense of involving judgements about clarity, similarity, analogy, and coherence. The alternative of phrasing these judgements in terms of the processes which generate

them tends to overemphasize the algorithmic aspect of rule formation and thereby obscure the all-important judgmental aspect.[16]

Furthermore, the reduction of judgements to the processes underlying them may render the argument unnecessarily prolix, possibly up to the point where the straightforward judgmental processes involved in rule formation become buried in a heap of conjectures about psychological processes which are largely irrelevant to rule formation. If we look at a mathematical theorem, for instance, we can undoubtedly identify its truth with the proof given for the theorem, and the rules that govern the relevant reasoning. Yet there are many different proofs conceivable for any given theorem, and we may conceive the truth thereof as independent of the proofing procedure, as all different proofs give the same result. There are many ways, for instance, to prove Pythagoras' theorem, both geometrically and algebraically. The truth of the theorem is independent of the particular proof chosen. When we apply the theorem, we suppose that it is true, without reference to any particular method of proof. To insist on reconsidering the proof over and over again would curb the usefulness of Phythagoras' theorem considerably. The theorem is useful because we can take it as given – without the underlying processes of proving it again and again.

In a similar vein, we may approach rule formation as a process which is driven by aesthetic judgement, without necessarily enlarging on how these aesthetic judgements themselves come about. In this sense, the aesthetic approach offers a shortcut which side-steps some issues in evolutionary psychology. The question of how aesthetic judgements are generated is largely irrelevant to the issue of rule formation. It suffices that these judgements are made, and are prompted by human psychological propensities that can be safely assumed as given and invariable in historical time.

This shortcut seems appropriate because the question about the formation of aesthetic judgement is fundamental, very difficult, and remains largely unresolved in evolutionary theory. Darwin himself emphasized the difficulty of accepting that 'mammals, birds, reptiles, and fish' share the 'high taste of beauty' which 'generally coincides with our own standard'.[17] Yet, according to him, aesthetic taste must be presupposed if we want to understand, for example, the phenomena of the peacock's tail-feathers or other significant features of animals in evolutionary terms. He invokes the idea that this sharing of aesthetic judgements across species may relate to the idea of common descent of all vertebrates, and that 'the nerve-cells of the brain in the highest, as well as in the lowest members of the Vertebrate series, are derived from those of the common progenitor of this great Kingdom.' The range of shared aesthetic judgements required for the present purpose is much more restricted, and less demanding, as it relates to humans only, and need not apply across species. In view of Darwin's observation on the role of beauty in evolution it would, however, be entirely mistaken to reject the relevance of aesthetic judgements in social co-ordination on evolutionary grounds. The argument that we know very little about the inner mechanisms of the aesthetic

16 This would be, in the terminology of Kubon-Gilke/Schlicht-1993, pp. 259-60, a 'conceptual implication' of the structuralist parlance.

17 Darwin-1874, p. 640.

sense does not imply that aesthetic judgements are irrelevant to biological and social evolution. Quite to the contrary: The fact that aesthetic judgements are made and widely shared offers a prima facie reason for assuming that they are evolutionarily significant. The importance of aesthetic judgement in learning processes offers further, an avenue of thought which may help us to understand what Darwin took as a fact: That we are endowed with an aesthetic sense.

9. Selecting for Learning

From a biologist's point of view, learning is interpreted as adaptive responses brought about by selective pressure. Learning is just a special case of adaptation. It involves the gathering and transmission of information. In this sense, evolution is a process of learning. We may envisage different levels of adaptation: the genetic, the individual, and the social level. Let us consider these in turn.[18]

1. Genetic learning. The process of biological evolution is typically envisaged as brought about by variation and selection. Genetic mutation and recombination generate variation. While well adapted individuals survive and multiply, the less well adapted are pruned off in the struggle for survival. Furthermore, the speed and direction of mutations is controlled by genetic mechanisms, which have evolved in the same manner. This gives rise to directed or patterned, rather than random mutation.[19]

2. Individual learning. However, not all organisms function like genetically programmed automata. In changing environments, genetic adaptation is sometimes too slow to track change. So some species have acquired the ability to learn and thereby to adapt more quickly.[20] This type of learning depends on recognizing

[18] Selten-1991, Jablonka/Lachmann/Lamb-1992, Lachmann/Jablonka-1996, and Jablonka/Lamb/Avital-1998 inspire the considerations in this section. Selten-1991, p. 21 distinguishes mutation, changes in gene frequencies (which I lump together), cultural transmission, and individual learning and stresses the different time dimensions involved. Jablonka, Lamb, and Avital distinguish, however, four inheritance systems: the epigenetic inheritance system, the genetic inheritance system, the behavioral inheritance system, and the linguistic inheritance system. The above classification amalgamates their epigentic and genetic inheritance systems. Further, as I am interested not only in inheritance, but more generally in learning, I distinguish here individual learning and social learning, which replaces their behavioral and linguistic systems to some measure. The fundamental argument introduced by Jablonka *et al.*, namely, that the different systems have their particular advantages under different conditions and will be selected for accordingly, is maintained.

[19] 'Some genetic structures do not adapt the organism to its environment. Instead, they have evolved to promote and direct the process of evolution. They function to enhance the capacity of the species to evolve.' (Campbell-1985, p. 137). Thus, evolutionary processes of variation must be assumed to be structured and patterned, rather than random and diffuse, see Jablonka/Lamb-1995: Chs. 3-5 and, with respect to social theory, Schlicht-1997.

[20] By the way, this observation puts into question a central tenet of evolutionary psychology, namely that evolution would favor domain-specific rather than general solutions in learning. The argument is that task-specific optimization is better at each task than any general strategy which could be applied to many tasks. (The issue of act-utilitarianism versus rule-utilitarianism re-appears here in a different guise.) The counter argument is that repeated task must be expected being automated and even genetically assimilated anyway. The *raison d'être* of learning is, thus, to cope with new issues in the best possible way, but there will be no chance for full optimization. See Shapiro/Epstein-1998 for related discussion.

recurrent patterns, identifying similar cases and forming hypotheses in the most efficient way. Learning relies on the 'supposition that the future resembles the past'.[21] Yet our ideas of resemblance must be prompted by correlations in the environment.[22] They cannot be fine-tuned to any one particular case because they constantly have to deal with new ones. This type of learning has proved successful, and has evolved, just as directed variation has superseded random mutation at the genetic level for evolutionary reasons.

3. Social learning. Genetic adaptation can be expected to occur in environments that remain invariant over time. Learning at an individual level can be expected to occur within environments that incessantly present new challenges to the individual. On an intermediate time-scale we can imagine changes which can neither be tracked by genetic change nor by individual learning in any satisfactory way. Let us envisage changes that occur over approximately a hundred generations. This time-span is too short to allow for significant genetic adaptation, but long enough to make it would be a waste of resources if each individual had to learn anew about the environment. Under such circumstances, it is more efficient for the individual simply to copy the behavior of others, rather than to find out about the environment on his own. This is when social learning evolves, and social tradition forms. In this social context learning relies on pattern recognition. However, the individual will be concerned with detecting patterns in the behavior of its conspecifics, rather than learning about the natural environment directly, which would be more costly. Once the customary behavioral patterns are assimilated, the individual may, through individual learning, improve on them and transmit improved behaviors to the next generation. This process gives rise to social evolution.[23]

10. Selecting for Rule Preference

Learning, whether social or individual, is concerned with recognizing regularities and recurrent patterns. These patterns, once recognized, help in guiding the individual's future behavior and eliminating that which is likely to fail. Learning prevents certain behaviors from being tried out. This strategy is certainly not the best, as it would be better to select the optimum solution in each specific case, but this is unrealistic; otherwise genetic encoding would have succeeded in producing such a response by now. The importance of an aesthetic sense to

21 Hume-1740, p. 134.

22 This is the theme in Lorenz-1973.

23 See Cavalli-Sforza/Feldman-1981, Boyd/Richerson-1985. As the theoretical argument suggests, processes of social learning and social evolution are not restricted to humans, but widespread in the animal kingdom, and give rise to a host of animal cultures and animal traditions; see Avital and Jablonka (in preparation).

enabling learning offers an argument for why we find individuals endowed with aesthetic preferences.[24] But we can go further.

Learning relates to novelty, and to detecting newly occurring patterns. In order to detect these, the individual must be *interested* in finding such patterns. Without an active interest in observing resemblances, analogies, and regularities spanning certain categories, they would go undetected. Take two individuals: One is interested in finding patterns, the other is not. In every other respect, both individuals are absolutely identical. Assume further that the environment is such that fitness can be increased by learning, either because it enables individuals to benefit by assimilating the knowledge encoded in the culture they live in, or by exploiting some idiosyncratic features of their particular habitat more effectively. Under these circumstances we must assume that an active desire for pattern recognition will increase fitness. The more curious individual – the one who likes and enjoys detecting patterns, similarities, and analogies – will be more successful than the disinterested one. In this way, we must assume natural selection to mould a sense of beauty, and an active desire to uncover patterns, just as we are endowed with a preference for nutritious food.[25]

As an aside, let me note that many inorganic things strike us as beautiful: crystals, rocks, a rainbow, the shapes of clouds, a waterfall in the sun. That we perceive these structures as beautiful indicates that our sense of beauty is tuned to such things, and there is a selective value in having such a taste. Furthermore, many aspects of beauty in animals, like the leopard's spots, have been traced back to the nature of physical and chemical processes, which severely channel and constrain both natural and sexual selection.[26] This strengthens, again, the point that aesthetic judgements are not arbitrary but reflect the structure of the universe in a deep sense for reasons we cannot easily understand.

This is highlighted also by the observation that the power of aesthetic judgements in uncovering the laws of nature is absolutely stunning. The physicist Paul Dirac was prompted by aesthetic reasons to reformulate an equation for the electron, which then led to the successful prediction of antimatter. He thought that 'it is more important to have beauty in one's equations than to have them fit the experiment.'[27] In a similar vein, the physicist Roger Penrose holds that 'rigorous

[24] Proponents of focal point arguments, like Schelling-1969 and Sugden-1986 rely in this sense on aesthetic judgement, but do not relate this to an aesthetic preference which is central for my own theory (Schlicht-1998).

[25] This argument has its limits, because curiosity and playfulness come at the cost of wasting time. We may, thus, postulate that evolution has settled for an appropriate level of such endeavors. Further, the above argument assumes that the desire to uncover and enjoy patterns is what we call the sense of beauty.

[26] Goodwin-1994. Note that these arguments differ: The fact that inorganic patterns strike us as beautiful can be interpreted in two different ways. One possibility (emphasized by Lorenz-1973) is that our aesthetic judgement has evolved because the physical world has properties that selected for correspondence between perception, cognition, and aspects of non-organic reality. The other possibility is that internal constraints such as those emerging from the way our nervous system is organized, provide the anchor for our aesthetic sense. For our present purposes we need not opt for one alternative or the other; and both may interact.

[27] Davies-1992, p. 176.

argument is usually the *last* step! Before that, one has to make many guesses, and for these, aesthetic convictions are enormously important.'[28] It has been noted that 'there is something curious here. If beauty is entirely biologically programmed, selected for survival value alone, it is all the more surprising to see it re-emerge in the esoteric world of fundamental physics, which has no direct connection with biology. On the other hand, if beauty is more than mere biology at work, if our aesthetic appreciation stems from contact with something firmer and more pervasive, then it is surely a fact of major significance that the fundamental laws of the universe reflect this "something".'[29]

This was just an aside to illustrate the astonishing power of aesthetic considerations in theory formation. We are not concerned here with these deep issues, but rather with everyday learning phenomena which build, however, on the same tendencies of thinking which guide the physicist in solving the riddles of the universe.

11. Conclusion

The point made in this paper is that aesthetic judgements and an associated active desire to uncover, maintain, and expand regularities is the source of rule formation in social interaction. The argument can be briefly restated as follows: All learning and extrapolation presupposes aesthetic judgements concerning similarity, analogy, simplicity, and straightforwardness. Learning has evolved as a response to changing environments, where genetic adaptation is too slow. It is,

[28] Davies-1992, p. 177. I must be added here that Einstein placed great emphasis on the 'truly religious conviction that this universe of ours is something perfect and susceptible to the rational striving for knowledge.' Here, 'perfection' cannot refer to purpose and must be, thus, taken as a judgement of an aesthetic kind. Einstein remarks that the search for perfection is of importance for the development of science: 'If this conviction had not been a strongly emotional one and if those searching for knowledge had not been inspired by Spinoza's *Amor Dei Intellectualis*, they would hardly have been capable of that untiring devotion which alone enables man to attain his greatest achievements.' (Einstein-1954, p. 52)

[29] Davies-1992, p. 176. Note, however, that Darwin took beauty not so much as biologically programmed, but rather as programming biological selection, and in particular sexual selection. This contrasts with modern treatments like Barrow's (Barrow-1995) which speculate that the human sense of beauty is shaped by the neolithic conditions our ancestors were exposed to. According to this argument, we like savannah-type landscapes because these provided the most comfortable environment for our Neolithic ancestors (Barrow-1995, p. 92; see also Richter-1999). Such arguments fall short of explaining why a polar landscape strikes us as beautiful and, more importantly, it does not address the universal aspects of beauty judgements which were Darwin's central concern.
The thought that learning presupposes an aesthetic sense may contribute to approach the issue in a Darwinian spirit. If the sense of beauty were fully adaptive, the pea-hen would prefer males with shorter tails for fitness reasons. This rules out the adaptive explanations mentioned above. If the aesthetic sense is shaped with respect to the efficacy of learning processes, however, it may entail those inefficiencies in sexual selection Darwin was concerned with (Darwin-1874). Our sense of beauty would then be adaptive with respect to learning, but would imply inefficiencies in other dimensions, like the peacock's tail.

however, not a passive phenomenon as it becomes particularly effective if the individual tries to actively uncover and exploit regularities in its environment. Hence evolutionary forces have instilled a rule preference – a desire to uncover, maintain, and expand patterns – as part and parcel of human nature. This rule preference gives rise to rule formation in social interaction. The argument provides, thus, an evolutionary underpinning for the 'clarity' view of custom.

References

Anderson, J. R. (1980) *Cognitive Psychology and its Implications,* third edition, New York: Freeman 1990.

Avital, E. & Jablonka, E. (to appear) *Animal Traditions,* Cambridge: Cambridge University Press.

Barrow, J. D. (1995) *The Artful Universe,* Oxford: Clarendon Press.

Binmore, K. & Samuelson, L. (1994) 'An Economist's Perspective on the Evolution of Norms,' *Journal of Institutional and Theoretical Economics,* Vol. 150, n°1, pp. 45-63.

Boyd, R. & Richerson, P.J. (1985) *Culture and the Evolutionary Process,* Chicago and London: The University of Chicago Press.

Buchanan, J.M. (1994) 'Choosing What to Choose', *Journal of Institutional and Theoretical Economics,* Vol. 150, n°1, pp. 123-35.

Buchanan, J. & Brennan, G. (1985) *The Reason of Rules,* New York: Cambridge University Press.

Campbell, J.H. (1985) 'An Organizational Interpretation of Evolution', in: *Evolution at the Crossroads: The New Biology and the New Philosophy of Science,* edited by D. J. Depew and B. H. Weber, Cambridge (Mass.): MIT Press, pp. 133-167.

Cavalli-Sforza, L.L. & Feldman, M.W. (1981) *Cultural Transmission and Evolution,* Princeton: Princeton University Press.

Darwin, C. (1874) *The Descent of Man,* New York: Crowell, reprint Amherst: Prometheus 1998, with an introduction by H. James Birx.

Davies, P. (1992) *The Mind of God,* New York: Touchstone.

Einstein, A. (1954) *Ideas and Opinions.* New York: Crown.

Goodman, N. (1983) *Fact, Fiction, and Forecast,* 3rd edition, Cambridge (Mass.): Harvard University Press.

Goodwin, B. (1994) *How the Leopard Changed its Spots,* New York: Scribner.

Goyal, S. & Janssen, M. (1996) 'Can We Rationally Learn to Coordinate?,' *Theory and Decision;* Vol. 40, n°1, pp. 29-49.

Hume, D. (1740) *A Treatise on Human Nature,* edited by L. A. Selby-Bigge, Oxford: Clarendon Press1888.

Hume, D. (1777) *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals,* ed. L. A. Selby-Bigge, second edition, Oxford: Clarendon Press 1902.

Jablonka, E. & Lamb, M. (1995) *Epigenetic Inheritance and Evolution. The Lamarckian Dimension,* Oxford: Oxford University Press.

Jablonka, E., Lamb, M.J. & Avital, E. (1998) '"Lamarckian" Mechanisms in Darwinian Evolution,' *Trends in Ecology and Evolution,* Vol. 13, n°5, pp. 206-210.

Jablonka, E., Lachmann, M. & Lamb, M.J. (1992) 'Evidence, Mechanisms, and Models for the Inheritance of Acquired Characters,' *Journal of Theoretical Biology,* n°158, pp. 245-268.

Kubon-Gilke, G., & Schlicht, E. (1993) 'Gefordertheit und institutionelle Analyse am Beispiel des Eigentums,' *Gestalt Theory,* Vol. 15, n°3/4, pp. 257-273.

Lachmann, M. & Jablonka, E. (1996) 'The Inheritance of Phenotypes: an Adaptation to Fluctuation Environments,' *Journal of Theoretical Biology,* n°181, pp. 1-9.

Lorenz, K. (1973) *Behind the Mirror,* trans. from German by R. Taylor, New York: Harcourt Brace Jovanowich 1977.

Maynard-Smith, J. (1978) "The Evolution of Behavior," *Scientific American,* n°239, pp. 176-192

North, D. (1990) *Institutions, Institutional Change and Economic Performance,* Cambridge: Cambridge University Press.

Pirker, R. & Rauchenschwandtner, H. (1998) 'Sense of Community: A Fundamental Concept of Institutional Economics,' *Journal of Institutional and Theoretical Economics,* Vol. 154, n°2, 406-421.

Richter, K. (1999) *Die Herkunft des Schönen,* Mainz: Philipp von Zabern.

Schelling, T.C. (1960) *The Strategy of Conflict,* Cambridge (Mass.): Harvard University Press.

Schlicht, E. (1997) 'Patterned Variation: The Role of Psychological Dispositions in Social and Economic Evolution,' *Journal of Institutional and Theoretical Economics,* Vol. 153, n°4, pp. 722-736.

Schlicht, E. (1998) *On Custom in the Economy,* Oxford: Clarendon Press. (An electronic version may still be available for free at the Oxford University Press Reading Room http://www1.oup.co.uk/academic/readingroom/schlicht/book.pdf.)

Selten, R. (1991) 'Evolution, Learning, and Economic Behavior,' *Games and Economic Behavior,* n°3, pp. 3-24.

Shapiro, L., & Epstein, W. (1998) 'Evolutionary Theory Meets Cognitive Psychology: A More Selective Perspective,' *Mind and Language* Vol. 13, n°1, pp. 171-194.

Smith, A. (1759) *The Theory of Moral Sentiments,* Oxford: Clarendon Press 1976.

Sugden, R. (1986) *The Economics of Rights, Co-operation and Welfare,* Oxford: Blackwell.