

"Tunnel effects" in cognition : A transfer mechanism from known conceptual domains to new ones

Antoine CORNUEJOLS¹ ; Andrée TIBERGHIE² ; Gérard COLLET²

¹ Laboratoire de Recherche en Informatique (LRI)
UMR 8623 CNRS
Université de Paris-sud, Orsay
Bâtiment 490, 91405 ORSAY (France)

² UMR GRIC-COAST
CNRS - Université Lyon2
5 Avenue Pierre Mendès France
69676 BRON Cedex 11 (France)

email : antoine@lri.fr ; Andree.Tiberghien@univ-lyon2.fr ; gcollet@ac-grenoble.fr

Abstract

This paper reports on a multidisciplinary head-on approach to the problem of learning new ways to interpret the world by relying on (and relating to) old ones. By studying how high school students address problems in conceptual domains that are new to them, we were led to analyze mechanisms that seemed to be at play in their segmenting the world, and constructing models of the situation, as well as the (re)conceptualization efforts that —sometimes— followed. In this paper, we focus on a reasoning mechanism that we hypothesize does explain part of the students behavior and may be at play within scientific discovery processes. Like analogy, this mechanism that we call tunnel effect by reference to transport effects that can take place in quantum physics, allows the transfer of knowledge from one conceptual domain to another one. Unlike analogy however, it does so without having to resort to two situations or cases, but only considers the one at hand, and it does not necessitate to specify beforehand a hierarchy of representation primitives in both domains (one being mostly unknown), nor to define how similarity between the two represented cases must be computed.

1 Introduction

Conceptual domains allow to disentangle the flow of perceptions from the world. They are tools that bring both a filter and a magnifying glass on our universe. They select and organize. They segment and predict. They allow to describe, to make predictions and to explain our environment. As such, when they are mature, they offer an operational way of tackling the world. Scientific discovery as well as education is concerned with learning conceptual domains. Surprisingly, in regard to their importance, few works have directly dealt with it. Philosophers have chosen to make a distinction between the context of creation and the context of justification of a theory, focusing almost exclusively on this last problem (see (Popper,1959;1962)). Cognitive scientists on the whole, and machine learning scientists in particular, have

mostly centered their work on induction of "simple" and isolated concepts in rather poorly structured hypotheses spaces. While the results obtained have been spectacular, in many respects they do not bear on the learning of complex conceptual domains and theories. In addition, the relatively few relevant works in machine learning (under the name of constructive induction, theory revision or inductive logic programming) have so far shared a common and mostly tacit assumption in that when learning a conceptual domain, the existing ontology of concepts was supposed to be correct, even if not always operationally efficient. The problem was thus seen as the one of learning new concepts *besides* existing ones, for instance by learning new concepts or predicates within the ontology in order to make it more efficient to use or more easy to understand. In this respect, the problem of learning a *new* conceptual domain was not really touched upon. In contrast, we see, in science and education, that one vital

problem is to learn new concepts and new ontologies, at once articulated with past ones, but also in competition with them.

That learning a new domain often implies that parts of the existing ontologies will have to be inactivated or modified according to the context is what we see as the most significant novelty of the problem we present here, and one that deserves much more attention than given so far.

This paper reports on a multidisciplinary head-on approach to the problem of learning new ways to interpret the world by relying on (and relating to) old ones. By studying how high school students address problems in conceptual domains that are new to them, we were led to analyze mechanisms that seemed to be at play in their segmenting the world, and constructing models of the situation, as well as the (re)conceptualization efforts that —sometimes— followed. In this paper, we focus on a reasoning mechanism that we hypothesize does explain part of the students behavior. We call it 'tunnel effect' for reasons that will be clarified later on. Like analogy, this mechanism allows the transfer of knowledge from one conceptual domain to another one. Unlike analogy however, it does so without having to resort to two situations or cases, but only considers the one at hand, and it does not necessitate to specify beforehand a hierarchy of representation primitives in both domains (one being mostly unknown), nor to define how similarity between the two represented cases must be computed. In fact, it appears so natural that its scope covers a wide range of situations from metaphorical thinking to scientific discovery (See for other descriptions (Nersessian, 1992; Thagard, 1992)).

In the following, we first describe, in section 2, a simple but telltale experiment in physics teaching. Section 3 describes how the mechanism of tunnel effect could account for it. How learning can occur as a result of tunnel effects is discussed in section 4, while section 5 contrasts tunnel effect with analogical reasoning with respect to the transfer of knowledge from one conceptual domain to another. Section 6 concludes by underlying key ideas and perspectives.

2 Illustration of the 'tunnel effect' in physics teaching

In order to study learning of new conceptual domains, we set up interpretation tasks in terms of a “new theory”. The idea was to force natural cognitive agents

to learn a new way to interpret the world, and to study how they tend to do it. More specifically, we performed experiments in physics teaching, and more precisely teaching a qualitative account of the physics of energy taught in high school classes around the age 16-17. The task involved small experimental settings that the students could experiment with, like simple electrical circuits with masses and motors and so on, that were to be interpreted in terms of energy transfers and transformations along an “energy chain” starting and ending with an energy reservoir. The students worked in pairs¹. This experiment has been done in several classes. We video-recorded several pairs of students and entirely transcribed their verbal productions. (For this paper 7 pairs were deeply analyzed).

On one hand, it is important to notice that the interpretation task was not trivial, even in the simplest of the experimental settings shown in figure 1. For instance, there were two wires from the battery to the bulb which satisfied the closed electrical circuit condition, but only one counterpart, standing for the transfer of energy under the form of electrical work, in the target interpretation. Likewise, the students had to discover the environment entity while there was no concrete, tangible, counterpart in the experimental setting.

On the other hand, the task facing the students was easier than the one facing the scientists in that they did not have to “invent” the concepts necessary for the task. They were indeed provided beforehand with a declarative account of the target conceptual domain along with a lexicon of the authorized terms and icons that were to be used in their models of the situation (see figure 2). The seed target domain also defined integrity rules that specified valid models, as, for instance, the “*a complete energy chain starts and ends with a reservoir*” rule. Together, the lexical entities used in the definition of the seed conceptual domain and the integrity rules constitute the target constraints for this particular task.

¹ In fact the students are given successively three tasks, only the first task is discussed in the paper. In the first task the experimental material is made up of a bulb, two wires, a battery. In the second task the experiment consists of an object hanging on a string which is completely rolled round the axle of a motor (working as a generator). A bulb is connected to the terminals of the motor. When the object is falling, the bulb shines (figure 3). In the third task the experiment consists in a battery connected to an electrical motor. An object is hanging from a string, attached to the axle of the motor, which is completely unrolled at the beginning. A correct solution is given to the students after the first task.

Now, we invite the reader to try for a few minutes to think of a program that could solve the problem above and others of the same type. Beware that the original description of the experimental setting is in itself a tricky problem. Some students for instance paid attention to details not shown here, like the electrical switch, the fingers, the eyes. Almost none however “perceived” the environment as an entity. While all of them treated the two wires as two distinct entities, most

1st Experiment

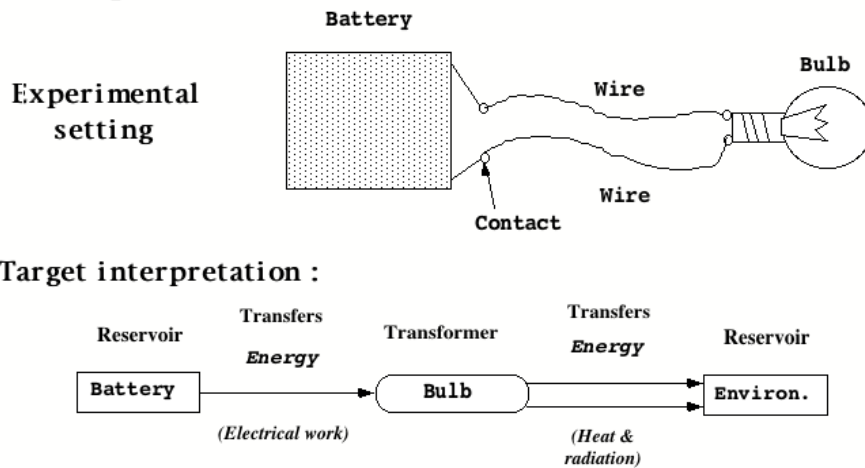


Figure 1. Above : one experimental setting involving a battery connected to a luminous bulb through two wires. Students were to produce an interpretation of this setting in terms of a chain of energy transfers and transformations starting and ending with an energy reservoir. Below : a correct interpretation, called target interpretation.

Theory (seed)	Model (seed)
Energy can be characterized by: * its properties : - <i>Storage</i> - <i>Transformation</i> - <i>Transfer</i> - by <i>work</i> : mechanical or electrical - by <i>heat</i> , - by <i>radiation</i> . * a fundamental principle of conservation The energy is conserved whatever the transformations, transfer and forms of storage	* <i>Symbols</i> to be used: <div style="display: flex; align-items: center; margin-bottom: 10px;"> <div style="border: 1px solid black; padding: 2px 10px; margin-right: 10px;">r es.</div> for reservoir </div> <div style="display: flex; align-items: center; margin-bottom: 10px;"> <div style="border-top: 1px solid black; border-bottom: 1px solid black; width: 20px; margin-right: 5px;"></div> <div style="border-left: 1px solid black; border-right: 1px solid black; width: 10px; height: 10px; margin-right: 5px; display: flex; align-items: center; justify-content: center;">▶</div> for transfer </div> <div style="display: flex; align-items: center; margin-bottom: 10px;"> <div style="border: 1px solid black; border-radius: 50%; padding: 5px 15px; margin-right: 10px;">tr.</div> for transformer </div> * Under the <i>constraints</i> : - a complete energy chain starts and ends with a reservoir; - the initial reservoir is different from the final reservoir.

Figure 2. A simplified version of the seed for the target conceptual domain given to the students. The left part presents the conceptual definitions for the target domain . The right part provides the symbols with which to express the model and the syntactic rules that should be satisfied.

did not single out the filament inside the electrical bulb. All in all, even in this extremely simplified setting, *the perception and interpretation of the experiment involve an incredibly large collection of choices, both local and low level and global and strategic.* Is there any way that this can be made otherwise ? Is there any way to help solve problems in an as yet ill-mastered domain ?

One fact that emerged from our study was that out of 7 pairs of students, 6 produced the intermediate model of figure 4 (b) below for the battery-bulb setting. They then departed from it to try to find alternatives, better suited models, meantime laboring over concepts like *energy*, *transfers*, and so on. This, in fact, did not strike us as worth of interest at first, so much it appeared to be expected. This intermediate model was after all none other than the classical circular electrical interpretation

of the setting. Yet, upon reexamination, we were intrigued by the fact that this model, which acted as a powerful attractor, seemed also pivotal to enable further conceptual elaboration. Did the analysis of the why and how of this particular behavior could lead to a better understanding of the processes at play in the learning of new conceptual domains? The rest of the paper is an answer to this.

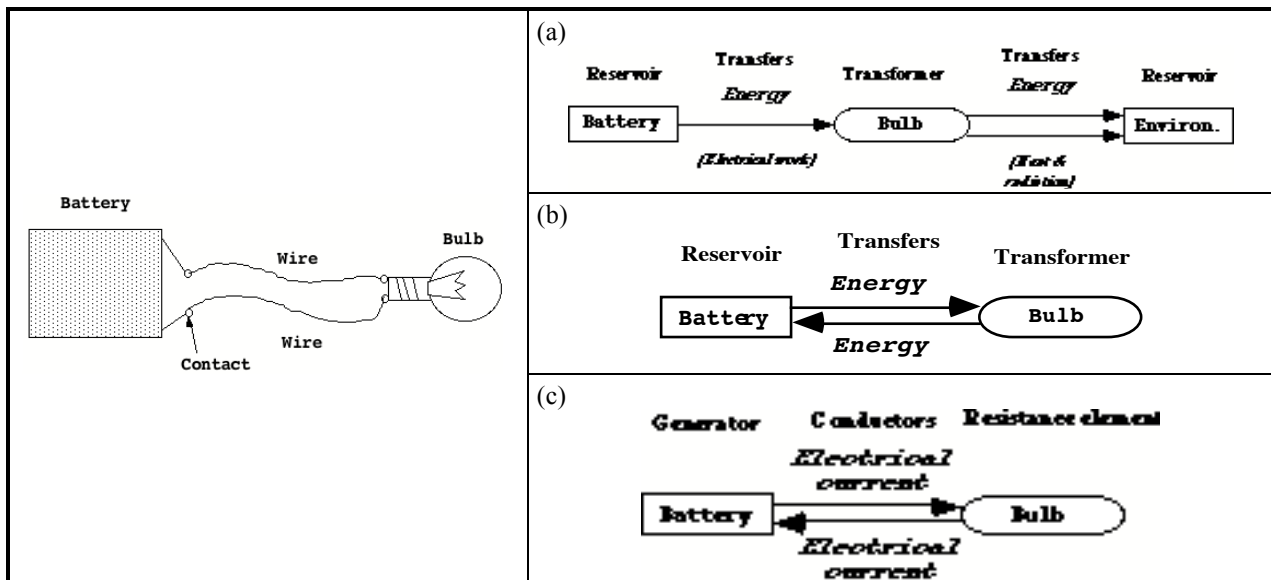


Figure 4. Three interpretations of the experimental setting of the left column.

3 The tunnel effect as an inferencing mechanism

In this section, we start by defining what we mean by conceptual domain. We will then emphasize several properties that concur in triggering tunnel effects before describing their overall mechanism.

I- We want to insist here on two attributes of *conceptual domains*. First, when mature, a conceptual domain can function entirely as a closed system with entities entertaining relationships with other entities of the same domain and defined only within this domain. When looked that way, that is ignoring its semantics, a conceptual domain can be characterized by some kind of *meta constraints* that specify the rules for well-formed formulas and acceptable derivations. The seed theory for energy chains is an instance of such meta constraints. Second, for the conceptual domain to be viable, this requires that it possesses a good *adequacy to the world*, that is that its predictions are reasonably confirmed and that it allows coherent and sufficiently complete description of the world (within the limits of

the implicitly defined viewpoint). We call *target constraints* these two requirements (meta constraints and adequacy to the world) in order to underline that a new, in construction, conceptual domain can thus be defined a priori.

For instance, while studying black body radiation, Planck was of course concerned that his theory would fit the experimental data (criterion of adequacy to the world), but also that it would give a picture of the world, which for him was deeply related to continuity, allowing to understand it (in particular how irreversible processes follow from conservative forces). He equally sets to himself that the theory should ensue only from the two first principles of thermodynamics (meta constraints). The drama for Planck was that he had to abandon the continuity criterion in order to fulfill adequacy to the experimental data and sufficiency of the two first principles of thermodynamics (a criterion for which he was ready "to sacrifice every one of (his) previous convictions about physical laws" (Planck, 1931)). However, in this case, we see how potent were the set of constraints deemed to be satisfied, and how they even

forced a completely new vision of the physical world, not wished for at first, whereby *quantum* physics followed.

Adequacy to the world and meta constraints as defined above are therefore enough to specify target conceptual domains *in a normative way*. They also provide means to judge the validity of new models or interpretations of some phenomenon.

II- Because here we consider *conceptual entities* that are part of interpretative systems, these entities are active and *carry with them expectations and inferencing mechanisms* not unlike the *if-needed*, *if-added* and other 'demons' that come with slots in schankian style conceptual primitives ((Minsky, 1975; Schank, 1982; Dyer, 1983). These mechanisms are responsible for prediction and completion of interpretation. They are also necessary in any interpretative systems which must construct a picture of the world from incomplete data.

III- It is generally the case that *new concepts have first to be described or thought of in terms of existing entities*, from other interpretative domains, that are irresistibly activated within the current context. For instance, students tend to first think of energy reservoirs as reservoirs with closures like most reservoirs of the everyday life. Only later, while encountering the environment or some weight acting as energy reservoirs will they revise their initial conception. The same can be said of many scientific developments : the concept of heat was painfully freed from the one of caloric itself associated with hydraulic connotations, the concept of speed was first deeply tied to the concept of the force causing the movement, making difficult to think of frames of reference and changes of these, and so on (Viennot,1996]. Every new conceptual domain is learned in interaction with existing interpretative domains. New concepts can be defined in terms of previous ones, as when Einstein combined the inertial mass and the gravitational one to define a new concept of mass within the general relativity theory. They can also be mistaken as one and the same as some other concept from another conceptual domain, like when students associate energy transfer with electrical current or when Sadi Carnot adopted the caloric interpretation of heat.

IV- *When two entities are associated, even if taken from different conceptual domains, their associated expectations and inferencing mechanisms are made*

available to both members of the pair for use if needed². Thus, when students associate energy transfers with electrical currents, and when there is a need for the determination of the direction of energy transfers, the inferencing mechanisms associated with the direction aspect of electrical current automatically come into play, even though they are foreign in the energy domain. These inference procedures are not thought upon and pondered, but on the contrary, they are smuggled in without further immediate checking. Hence the circular nature of the model of figure 4(b).

It is important to realize that this phenomenon, which is central in what we call the tunnel effect in cognition, is ordinary. It happened when Carnot was equating the "caloric" with heat, and thereby introducing — smuggling in— its conservative property. It happened to Maxwell when he equated the ether (incompressible fluid) with a model for electromagnetic interactions, smuggling in the seeds for the difficulties faced in physics until Einstein's special relativity theory got rid of them (and of most of the smuggled in properties of ether). It happens all the time, and it happens unconsciously. This smuggling might turn out to be genial when it brings with it unexpected solutions to outstanding problems. It might also hinder further solution.

To sum up, each time entities from two different interpretation domains are matched, they can potentially bring with them in these associations further attached properties that are new to the other entity. And this can happen in both ways. For instance, we noted that energy transfers found themselves naturally endowed with directions as soon as energy was associated with electrical current. Likewise, in another task not presented here, one student matched reservoir with a weighting object, then to show that the weight could be filled up (!) by being lifted. An example of a property not to be found originally in the notion of weight (source domain), but really brought by the contextual match with reservoir (target domain).

V- Properties II, III and IV are responsible for *tunnel effects* : the *implicit transfer of properties and inferencing mechanisms from one conceptual domain to another*. This is due to the fact that when a concept is transferred from one domain to another, some of its properties are checked for conformity with the target constraints (point I above) (e.g. electrical current seems

² For lack of space, we do not analyze here the mechanism underlying these associations. One key hypothesis is the existence of a *notional level*. We report the interested reader to [Cornuéjols et al., submitted].

to be fluid, linked to causality, and commonly represented with arrows) while other go unnoticed, and therefore unchecked (e.g. the circular property of electrical current which is in contradiction with integrity rules of the target domain : the initial and final reservoirs should be different). It is essential to stress that tunnel effect implies the actual transfer from one – source– domain to the other –target– domain. Indeed, the “solution” found in that indirect way, through a lack of differentiation with some source domain(s), is actually re-interpreted within the target conceptual domain. This is the reason why the students may discover the inadequacy of their solution because, according to the interpretation of their model within the energy conceptual domain, the energy flows back to the battery/reservoir, something which, they know, goes against their previous knowledge and against the target constraints.

To sum up, the *mechanism for tunnel effect* implies the following steps :

1. Some association is made between entities from some active and operational source domain(s) and the would-be entities of the target domain.
2. While building the interpretation of the situation in terms of the target domain, expectations and inferences from the associated source entities are illegally being used to fill up the missing aspects of the target interpretation.
3. The model thus built is then interpreted entirely within the target domain, free of its underlying, and possibly murky, justifications.
4. In case of uncovered problem in step 3, some reconceptualization work may occur, possibly contributing to the setting-up of a true “correct” and operational target conceptual domain.)

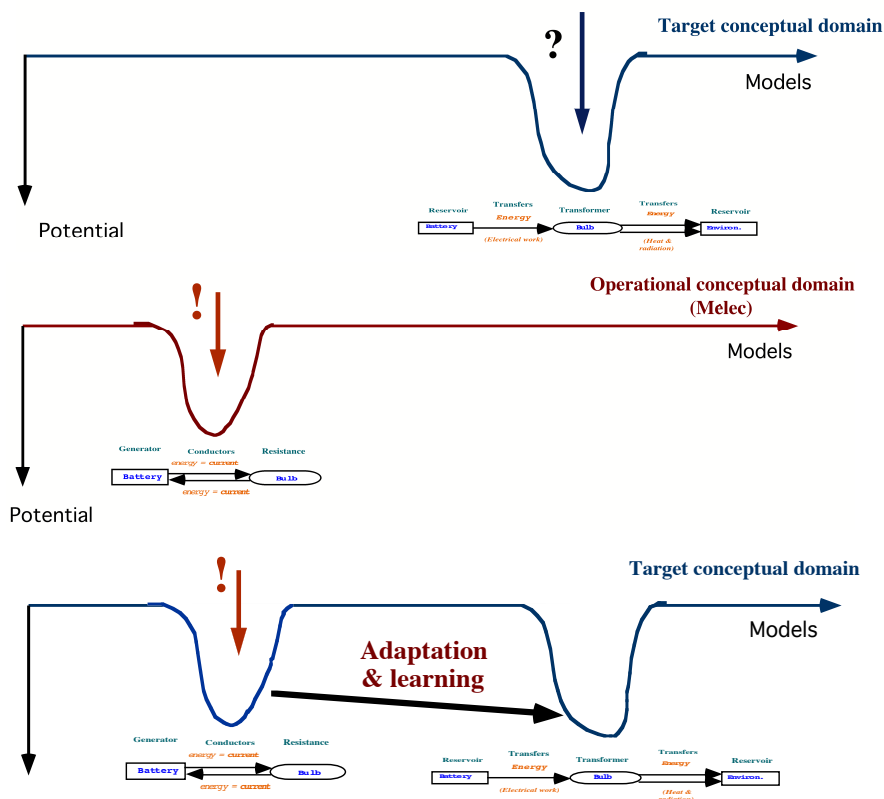


Figure 5. A decomposition process facilitating problem-solving in an ill-mastered conceptual domain. In each schema, the horizontal axis stands for the space of potential models and the vertical axis stands for the quality of the model with regards to the world. Of course, depending on the interpretation domain (for instance centered on electricity or on energy exchanges), the same models may have different degrees of quality.

3.3 Tunnel effect as a way to decompose problem solving

In response to the challenge to the reader in section 2, one can see tunnel effect as a way to ease problem-solving in an ill-mastered conceptual domain. For instance, no students were able to solve directly the first energy chain task. The problem was simply too hard for them. On the other hand, 12 out of 14 produced the intermediate model of figure 4(b), which is arguably an electric model of the setting disguised as a legitimate model in the energy domain. If it is difficult for an agent to solve directly a problem in an ill-mastered domain, it might be easier to disguise an interpretation stemming from well-known domains into a legitimate candidate model in the target domain. The question then is of course that of seeing if that step (a kind of forgery, except it may be unconscious resulting from automatic inferencing in the source domain(s)) helps or hinders further resolution of the problem.

It is difficult to answer this question in general, except that some version of the now famous *no-free lunch theorem* known in Machine Learning (Wolpert, 1992) and Optimization Theory is likely to apply and state that, overall, tunnel effect must equally ease and hinder problem-solving in new domains depending on the context.

However, there are reasons to think that tunnel effect may be a powerful help in problem-solving in some cases. Figure 5 suggests why. Thanks to tunnel effect, there are apparently more solutions to the interpretation problem, and hence more opportunities to find one of them. The problem then, if a fallacious solution has been found, is to be able to find a way towards the correct solution. We show in section 4 that this may be facilitated by the focus naturally provided by the processes underlying tunnel effect.

4 How tunnel effect activates further adaptation and conceptual learning

Two cases must be examined with respect to the opportunities for learning opened when a model has been obtained using tunnel effect :

1. The model obtained remains valid even after being re-interpreted in the target domain under construction (i.e. in figure 5, the holes in the different landscapes are superposed).
2. The model turns out to be erroneous either when confronted with the world or because internal

inconsistencies are discovered within the target interpretation domain.

4.1 The model remains valid

This is what happened during the construction of thermodynamics by Carnot, Clapeyron, Thomson, Joule, Clausius and others (Longair, 1984; Science & Vie, 1994). Carnot, influenced by the theory of the *caloric* (an imponderable fluid with the property of being conserved and which he equated to heat) and by his father's work on the calculation of the efficiency of water mills, devised a cyclic and reversible model describing an ideal steam engine. Thanks to this model, he was able to demonstrate that there exists a maximal efficiency for steam engines, and that it depends on the difference of temperature between the hot source of heat (caloric) and the cold one. Later on, through a series of very meticulous experiments, Joule was able to show that heat was not a conservative quantity and was exchangeable with work. However, it turned out that Carnot's model was in fact neutral with respect to the caloric hypothesis and when re-interpreted in the context of the new theory about heat and work, still remained a very helpful tool for thought experiments, one which eventually led to the discovery by Clausius of a special state function called entropy.

We have here one instance of a model obtained through tunnel effect (its cyclic and reversible nature was deeply a result of the belief in the caloric theory even though this was never explicitly expressed by Carnot) which is still valid once the interpretation domain changes. The model by itself cannot therefore act as a trigger for re-evaluation of the target domain, and other symptoms must show. However, because it remains valid, it can help shape the new conceptual system and serve as a test bed for it, potentially through thought experiments as this was the case for Carnot's model in thermodynamics.

4.2 The model turns out to be erroneous when re-interpreted

In our energy chain experiments, this happened either when students realized that the model implied that the energy was flowing back to the battery (which they knew was incorrect), or when they discovered an inconsistency with the target integrity rule stating that the initial energy reservoir should be different from the final one.

The natural question is then why is the model wrong in the investigated aspect ? A re-examination of the path that led to this conclusion in the model can then point

towards one of two causes. First, the associations made between entities from the target domain and the source one(s) could be erroneous. For instance, many students question the association they made between electrical current and energy or between the wires and the transfers. This can lead to a differentiation process whereby the target entities gain autonomy with respect to the source ones. Second, the automatic inferencing process that determined the problematic aspect of the model can be disclosed and limitations for its range been set. This is what happened when some students realized that the circular nature of the electrical current did not carry to the energy entity. This inference was henceforth stopped when building a model. (Cauzinille-Marmèche et al. (1997) provide an analysis of some “repair mechanisms” used by students to adapt their model).

This short discussion convincingly shows in our opinion that tunnel effects, not only help finding models, even erroneous ones, but that they also provide guidelines for further re-examination and reconceptualisation when needed. This is however an issue that deserves much further work.

5 The tunnel effect vs. analogical reasoning

Very few inference mechanisms have been proposed that deal with the transfer of information between different conceptual domains. *Analogical reasoning* is one of them —the most famous—, *blending* is another one (Fauconnier & Turner, 1998), and, we submit, *tunnel effect* is a contender too. A full comparative study of the three of them would be more than interesting, but is beyond the scope of this paper. However, we believe that a comparison with analogical reasoning might help to enlighten some characteristics of the tunnel effect as an inferencing mechanism. We will concentrate in each case on the conditions for a transfer between interpretation domains to occur, and on the information content that is transferred.

According to the dominant view on analogy (e.g. (Falkenheimer et al., 1989; Greiner, 1988)), **analogical reasoning** involves the interpretation of two cases, — called the source case for the supposedly well-known one, and the target case for the one to be completed—, that may be interpreted within two different interpretation domains (e.g. the solar system as a source case and the supposedly ill-understood atom system as a target one). Each case is supposed to be represented as a graph of relations and nodes standing for primitive concepts. Analogical reasoning implies then that a best

partial match be found between the two graphs, and, in a second step, that the part of the graph representing the source case with no counterpart in the target case representation be copied, translated and added to the target representation in order to fill the missing part. Many questions arise as to the principles that should govern both the matching operation, the translation and the transfer, not to speak about subsequent verification and adaptation. Deep concerns have also been expressed about the interpretation process of the two cases during analogy and the ensuing representation of the cases (e.g. (Hosftädter, 1995; Mitchell, 1993)). It is important to note that both domains —the source and target— must be sufficiently well understood in order that the respective conceptual primitives be identified, put in hierarchy and potentially matched. This view of analogical reasoning thus prevents the consideration of a target domain that would be in gestation and of which conceptual primitives would be very uncertain.

If we consider then the analogical inferencing mechanism as a kind of black box with inputs and outputs, the *inputs* consist in the source and target conceptual domains (the conceptual primitives and their relationships (including the said over-important hierarchies) and in the two cases (be they already represented as some would pretend is realistic, or be they interpreted in the context of the analogy as others would insist is unavoidable). The *black box* then searches for one satisfying matching between the two cases (given as rigid representations or not) and computes the completion of the target case representation. The *output or information gained* in the operation consists therefore in the added features and properties of the target case.

In contrast, **tunnel effect** only involves the interpretation of a *single situation or case* (e.g. an experimental setting or a set of phenomena). The *input* of the tunnel effect black box consists in the operational source interpretation domain(s), the target criteria that specifies the target interpretation domain (including preconceptions about some target entities, their properties and relationships), and the case (situation or set of phenomena) to be interpreted and understood in the target interpretation domain (e.g. the battery-lamp experiment to be interpreted in terms of energy exchanges, the electromagnetic interactions as measured in Faraday's experiments in terms of a theory in germ in Maxwell's head, or the steam engines in terms of heat and work and other related variables in the nascent thermodynamics). The *black box* then searches for a model of the case satisfying the target criteria. Because

most target entities are not yet operational and interpretable directly in the world, they have to be translated in terms of the more operational interpretation domains given as inputs. In this translation process, submitted to the target criteria, and during model building, some aspects of the model may be automatically filled up through automatic inferencing within the source domain(s) (as is the case when the arrows for transfers are automatically specified when it

is decided to translate energy transfer from the notion of electrical current). The *output or information gained* in the operation consists in the unexpected (because not planned) consequences of the model when interpreted within the target interpretation domain, or in the experimental setting if some target entities are already partially interpretable in the world (as is the case for "energy" for 16-17 years old students).

Table 1. A summary of the main features of analogical reasoning versus features of tunnel effect.

Analogy	Tunnel effect
<ul style="list-style-type: none"> • Two experimental settings or situations that are posited as analogs to each other • Interpretation takes place both in the source domain and in the target domain (there are two situations to be interpreted). • Relies heavily on comparisons : • Implies complex pattern matching between the two case representations • Tightly associated with the notion of similarity <i>between</i> structures. One problem is to explain how this similarity is computed • There is transfer by matching, alignment and completion from the source to the target • <i>New information</i> is produced through the completion of the target case representation • Does not explain how the source is chosen • Learning is supposed to arise as : <ul style="list-style-type: none"> - learning of indexing scheme - generalization and abstraction from analog cases - not really new conceptualization, except by generalization 	<ul style="list-style-type: none"> • One experimental setting or situation only • Interpretation takes place in the source domain subject to the target constraints and adequacy to the world criterion. • No comparison is involved, only interpretation • Involves associations at the notional level between target entities and source ones • Associated with confusion at the notional level. No notion of similarity <i>between</i> constructs • There is transfer by reinterpretation of the model of which some aspects have been automatically filled-in within the source interpretation domain(s). The built model gains autonomy and is reinterpreted in the target domain • <i>New information</i> is produced through automatic completion of the model within the source domain • The source domain(s) is(are) the most operational for interpretation in the current situation • Learning : <ul style="list-style-type: none"> - Reconceptualization focuses on associated entities that led to inconsistencies in order to differentiate them - Progressive operationalization of the new conceptual domain - Articulation with primitive perceptions about the world and with the source conceptual domain

In both analogical reasoning and tunnel effect, the detection of discrepancies between the resulting model and the world or of other inconsistencies opens opportunities for learning. The difference lies in the fact that tunnel effect is intrinsically intended towards the process of building the domain interpretation domain (through the setting up of connections between this

domain, the operational ones in the context and the world) whereas analogical reasoning is oriented towards the completion of some specific case with the help of another 'similar' one. While failed analogies may lead to reconceptualisation in the target interpretation domain, this is much less direct than the learning that may occur

when a tunnel effect has produced an unfit model of the world in the interpretation domain.

6 Conclusion

This paper takes seriously the idea that cognition may imply the existence (and coexistence) of several different interpretation universes, and that a specially important type of learning consists in acquiring new ways of interpreting the world or some aspects of it. In our study we focused on the passage from the currently operational interpretation domain(s) to a new target one when the attention of the cognitive agent is driven towards the interpretation and understanding of some phenomenon or set of phenomena.

When a new interpretation domain is learnt (i.e. new segmenting of the world and new inference rules), the new concepts and new rules are not yet settled nor directly interpretable in the world (think about the first time you heard of tensor calculus or of electrons). They have to be linked with known entities. Therefore, when a model is built in terms of target entities, it in fact refers to the world mostly through entities and relations belonging to the currently operational domain(s). Aspects of this model might thus be filled in thanks to automatic (and unchecked) inferences within the source domain(s). This is the basis for the tunnel effect. These added features, expressed in the model, when re-interpreted within the target domain may bring out unforeseen consequences.

Tunnel effect eases the construction of models by providing inference mechanisms from the source domain(s) that make up for the as yet non-existent inference mechanisms of the target domain. In so doing, erroneous models might be obtained. These intermediate models can help or hinder reaching a later, more adapted, model. Even though we think we have strong arguments in support that tunnel effect can be a powerful guide for further reconceptualisation (see section 3.4), this is still a matter for research, specially in view to the fact that, in case favorable conditions could be identified, one could envision using well-guided tunnel effects to ease the teaching of scientific domains and to apply them for machine discovery.

References

E. Cauzinille-Marmèche, G. Collet, A. Cornuéjols and A. Tiberghien. Co-adaptation of students' knowledge domains when interpreting a physical situation in

- terms of a new theory. In Proc. of the 2nd European Conf. on Cognitive Sciences (ECCS'97), Manchester, April 1997, pp.107-112.
- A. Cornuéjols, A. Tiberghien and G. Collet (submitted). Tunnel effects in cognition : A new mechanism for scientific discovery and education. Submitted to the *International Journal for Human-Computer Studies* (special issue on Machine Discovery).
- M. Dyer. *In-depth understanding*. MIT Press, 1983.
- B. Falkenhainer, K. Forbus, and D. Gentner. Structure-mapping engine, *Artificial Intelligence*, 41, 1-63, 1989.
- G. Fauconnier and M. Turner. Conceptual integration networks. *Cognitive Science*, vol.22 (2):133-187, 1998.
- R. Greiner. Learning by understanding analogies. *Artificial Intelligence* 35:81-125, 1988.
- D. Hofstadter. *Fluid Concepts and Creative Analogies*, Basic Books, 1995.
- M. Longair. *Theoretical Concepts in Physics*. Cambridge University Press, 1984.
- O. Megalakaki and A. Tiberghien. Corpus de dialogues de trois groupe d'élèves résolvant trois problèmes mettant en jeu une activité de modélisation, *CNRS-COAST Research Report*, no. CR-10/95, 1995.
- O. Megalakaki. Expérience MARIANNE : corpus de dialogues de trois gourpes d'élèves résolvant une séquence de problèmes mettant en jeu une activité de modélisation, *CNRS-COAST Research Report*, no. CR-11/95, 1995.
- M. Minsky. A framework for representing knowledge. In P.H. Winston (Ed.) *The psychology of computer vision*. McGraw-Hill, 211-277, 1975.
- M. Mitchell. *Analogy-Making as Perception*, MIT Press, 1993.
- N. Nersessian. How do scientists think? Capturing the dynamics of conceptual changes in science in Giere, R. (Ed.) *The Minesota Studies in the Philosophy of science*, vol.XV, University of Minnesota Press, Minneapolis, pp.3-44, 1992.
- M. Planck. Letter from M. Planck to R.W. Wood, 1931 (See Hermann, A., 1971, *The Genesis of Quantum Theory* (1899-1913), MIT Press, Cambridge Mass., pp.23-24).
- K. Popper. *The logic of scientific discovery*. Basic Books, NY, 1959.
- K. Popper. *Conjectures and refutations : The growth of scientific knowledge*. Basic Books, NY, 1962.
- R. Schank. *Dynamic memory*. MIT Press, 1982.

- E. Schrödinger. *Ma Conception du Monde*. (My vision of the world). Mercure de France, 1982, p.30.
- Science & Vie*. Special issue on Carnot and the discovery of thermodynamics (in French), 1994.
- P. Thagard. *Conceptual revolutions*. Princeton University Press, 1992.
- A. Tiberghien. Modelling as a basis for analysing teaching-learning situations. *Learning and Instruction*, 4(1):71-87, 1994.
- A. Tiberghien. Construction of prototypical situations in teaching the concept of energy. In G. Welford, J. Osborne, P. Scott (Eds) *Research in Science Education in Europe*. London: Falmer Press. p. 100-114, 1996.
- L. Viennot *Raisonnement en physique. La part du sens commun*. De Boeck Université. Paris-Bruxelles, 1996.
- D. Wolpert. On the connection between in-sample testing and generalization error. *Complex Systems*, 6, 47-94, 1992.