What, When and How do the Models of Conceptual Change Explain?

Anna-Mari Rusanen (anna-mari.rusanen@helsinki.fi)
ConChaMo/Department of Physics, PO BOX 64,
00014 University of Helsinki, Finland
Department of Philosophy, History, Art and Culture Studies,
PO BOX 24, 00014 University of Helsinki

Otto Lappi (otto.lappi@helsinki.fi)
Cognitive Science/ Institute of Behavioural Sciences,
PO BOX 9
00014 University of Helsinki, Finland

Abstract
There are certain theoretical issues in conceptual change research that are still puzzling researchers. First, there is no agreement on what kinds of changes in belief and concept systems constitute conceptual change. Second, there is no consensus on what the mechanisms of conceptual change are. Third, there is no common understanding how to explain, model and describe in an exact way these underlying mechanisms. In this paper, we offer a diagnostic analysis these issues by reviewing current theories of conceptual change in a framework of mechanistic explanations of cognitive phenomena, and present a possible sketch for explanations of conceptual change.

Keywords: Conceptual Change; Explanation; Cognitive Mechanisms

Introduction

Concepts enable thought, reasoning and problem solving. Acquiring new concepts or reorganizing the conceptual framework one uses to think about a domain is a powerful kind of learning. This sort of learning is known as conceptual change.

Conceptual change is one of the most studied fields in science education, and there are hundreds, if not thousands studies, on this topic. However, there are still some foundational issues in conceptual change research on which no clear consensus has emerged, and that are still puzzling most of the researchers. Firstly, there is no agreement on what kinds of changes in belief and concept systems actually constitute conceptual change. Secondly, there is no consensus on what the mechanisms of conceptual change are. Also, when compared to the level of detail at which, say, basic visual processing is understood, often the descriptions of these “mechanisms of conceptual change” are quite shallow and offer no exact specification of the precise structure of mechanisms.

As Clement (2008) remarks, there are very few, if any, models of conceptual change, in which the mechanisms of conceptual change are specified in sufficient detail. This suggests that many of the current accounts might not in a strict sense qualify as sufficient for explanation and manipulation of learning phenomena involving conceptual change.

Moreover, having numerous loose filler terms in an explanation does not only threaten to undermine its explanatory power, but filler terms may also be barriers to scientific progress when they veil failures of understanding (Craver, 2006, 2007). If, for example, the terms “reassign” or “assimilation” are used to stand for processes with largely unknown properties, then we really do not explain what happens. Instead, we have a possible sketch for an explanation. If this sketch is taken to be genuinely explanatory, then - in the worst case scenario - it is possible that we have only an illusion of explanation instead of having a genuine one (Rosenblitz & Keil, 2002; also Craver, 2006).

In what follows, we analyse the explanations of conceptual change from a philosophical point of view. Our analysis is partially based on the so-called “mechanistic account of explanation” (Bechtel & Richardson, 1993; Machamer et al., 2000; Craver, 2006,2007). This mechanistic approach has not previously been applied to explanations of conceptual change.

However, it should be emphasized that the focus of this paper is only on cognitive accounts of conceptual change. There are other accounts of conceptual change that examine conceptual change from socio-cultural, emotional or motivational perspectives. However, explanation of cognitive phenomena is a unique form of explanation, and it is an open question, whether it is possible to extend this form of explanation to cover explanation involving interpersonal dynamics etc. This topic is, however, beyond the scope of this paper.

Variable Accounts of Conceptual Change

The study of conceptual change has focused, on the one hand, on the acquisition of commonsense concepts in childhood (e.g. Carey, Spelke) and on the other hand, the
...acquisition of scientific concepts in science education, especially at the secondary and tertiary level (Chi, 1992; Chi et al, 1994; Vosniadou, 1992; DiSessa, 1993).

In this paper we examine the latter form of research, in which conceptual change is seen as a specific kind of learning process, in which a student does not merely accumulate more knowledge, but her conceptions of phenomena in a certain domain undergo a restructuring process that affects ontological commitments, inferential relations, and standards of explanation (Posner et al, 1982; Carey, 1985; Chi, 1992; Vosniadou, 1994; DiSessa, 1993). In a nutshell, this sort of conceptual change can be characterized as transformation process of the initial knowledge-state (a commonsense picture of the world) to one of various outcome knowledge states. The outcome state can be an accurate scientific conception (when the learning process has been successful) or, when the learning process has not been successful, one of a number misconceptions (when it has not).

One difficulty with conceptual change research is that there are a huge number of different accounts of the details conceptual change, and they all characterize conceptual change different ways. In the literature, there are different views on the learner’s initial and outcome conceptions, on the trajectories along which change occurs, on the mechanisms that are underlying the conceptual change, on the obstacles of learning and also on the factors that support the change. However, on the basis of a careful reading, the literature seems to suggest that there are roughly three “major” kinds of conceptual change. These different kinds, or forms, of conceptual change can be titled as revision, reinterpretation and invention:

Revision. In some cases conceptual change seem to require a revision of an existing conceptual system. For example, Chi and her colleagues suggest that conceptual change takes the form of category shift (Chi, 1992, 2008; Chi et al, 1994). Another example of this form of conceptual change is described by Thagard’s “tree jumping”, in which conceptual change happens when hierarchies of concepts are reorganized by shifting a concept from one branch of a hierarchical tree of concepts to another. Also in DiSessa’s Knowledge in Pieces account conceptual change is understood as a form of revisionary process, because conceptual change is seen as a process that integrates initially piecemeal, incoherent (sub)conceptual system by complex process of organizing and reorganizing the elements of the system (DiSessa, 1993, 2002, 2004).

Reinterpretation. In some cases conceptual change requires that a learner gives a new interpretation for a domain. For example, according to Ohlsson (2009) conceptual change occurs when a learner uses analogical transfer to map conceptual system from one domain to new domain to which it has not previously been applied, and to which some other conceptual system had been predominant. Another examples of conceptual change as a form of reinterpretation are described by Carey’s differentiations (when initially undifferentiated concept is differentiated) and coalescence (when initially distinct concepts are subsumed by a same concept).

Invention. In some cases conceptual change requires construction or production of a novel (for the learner) conceptual system. For example, Carey (1985, 2011) describes a form of conceptual change, in which a learner constructs a new set of concepts from already existing concepts by “bootstrapping” in way that makes novel concepts incommensurable with the earlier concepts, because the content of new concepts cannot be captured in terms of any previously possessed concepts. The first stage of bootstrapping, or “Quinian bootstrapping”, occurs when a learner encounters a set of explicit public symbols, such as sentences of a scientific theory. These public symbols, “the placeholders”, are not initially mapped onto any already existing concepts that a learner holds. Rather, for a learner they are either partially or completely uninterpreted. During the process of learning, these placeholders are then taken up by various “modeling processes”, which includes abstract forms of theoretical inference such as analogical reasoning, abduction and induction etc. The idea is that a learner constructs the interpretation or the content for a placeholder by using these different mechanisms. At the end of the process, the placeholders will have conceptual content in virtue of acquiring a stable conceptual role in a new theoretical structure.

Conceptual change as an umbrella term. This variety of different kinds of conceptual change might reasonably be taken to indicate that “conceptual change” is a sort of umbrella term, which covers several types of phenomena instead of referring to a singular type of learning. This would entail that there cannot be a singular “grand theory” of conceptual change, which could explain all possible instances and trajectories of conceptual change.

Instead, explaining conceptual change seems to require that different learning trajectories are explained by referring to different mix of underlying mechanisms and processes (see also Chinn & Samarapungavan, 2009; Ohlsson, 2009b). These learning trajectories can be considered conceptual change because the learning is seen in some way “radical”. However, this does not necessarily indicate that these different phenomena are instances of a common explanandum i.e. a common learning phenomenon, which the various mechanisms would account for. Instead, if one used this as a reason for adopting the umbrella term, it would be merely a pragmatic reason.

Towards the Explanation of Conceptual Change

The explanans and explanandum. In the case of conceptual change research, it is not always apparent, what the explanandum (the thing to be explained) and what the explanans (the things that explain) are. For example, Mayer...
Dissection of Explanation.

From a philosophical point of view, explanation of cognitive phenomenon typically involves at least (1) the characterization of the specific cognitive task performed by a system, (2) the description of how certain cognitive mechanisms execute, produce or sustain the phenomenon to be explained. In some cases, explanation in cognitive science also requires the (3) description of how cognitive mechanisms are implemented in cognitive systems or why it makes rational (or evolutionary) sense that the phenomenon should be sustained in the first place.

Explanation step 1: The characterization of the specific cognitive task. In cognitive explanations of behavioral phenomena, the description for the task is given by characterizing the information processing task, and it answers to questions such as: “What is the cognitive goal of this process?” or “What is the cognitive task of this competence?”

This aspect of explanation is important for two reasons; it not only characterizes the cognitive task in a specific way, but it creates also some constraints for the possible underlying mechanisms. This aspect of explanation characterizes, why certain - but not all - learning mechanism are appropriate for fulfilling the cognitive task.

The task of conceptual change. So, what is the task of conceptual change? Even if the issue of the task is not often expressed explicitly in current literature, many seem to echo the same normative intuition that the task of conceptual change is to reorganize the conceptual system in a way that makes - in a case of successful learning - somehow “better”. Depending on the larger picture of conceptual change, different authors have described this “better” different ways.

One early formulation can be found in the seminal paper by Posner et al. (1982), where they propose that conceptual change makes the system “more fruitful, intelligible and plausible” etc. (Posner et al, 1982). In their paper, intelligible means roughly that the new conception must be clear enough to make sense to the learner. Plausible means the new conception must be seen as believable, and even true. Fruitful means the new conception must appear potentially productive to the learner for solving problems and seek for new intellectual directions. The approach Posner et al. propose is based on the Kuhnian idea of paradigm shifts and their emphasis of the conceptual ecology of a student. By conceptual ecology Posner and colleagues meant the framework of a learner’s conceptions and “cognitive mechanisms”, such as analogies, metaphors, explanatory anomalies and so on (Posner et al. 1982; Strike & Posner, 1992). So, according to this view, conceptual change happens if the changes make the ecology “better” i.e. more productive and fruitful, and it increases the ability to solve problems.

Sometimes this “better” is interpreted in terms of utility. For example, Stellan Ohlsson recently proposed that some forms of conceptual change make the conceptual system more useful (Ohlsson, 2009). In Ohlsson’s account cognitive utility measures the usefulness of a knowledge system for a learner. The basic idea is that in a situation, where there are competing knowledge systems, the system that requires less cognitive load, and leads to faster, more efficient and more cognitively satisfactory end states, will become associated with higher strength and will be easier to activate (Ohlsson, 2009). Over time, the system will become the person’s “standard way of looking at the target domain” (Ohlsson, 2009).

In some cases, the task of conceptual change is given also in terms of coherence. For instance diSessa (1993, 2002, 2004) describes novice knowledge as a weakly organized system that is highly context dependent and internally inconsistent, thereby lacking internal coherence. In diSessa’s account commonsense physical knowledge is organized into p-prim, empirical typologies or low-level abstractions of everyday experience. For example, according to this knowledge-in-pieces- account novices’ knowledge systems are fragmented and consist of loosely connected pieces, which often lacks not only coherence but are also employed with little co-ordination (diSessa 1993, 2002, 2004). In diSessa’s and colleagues account, the task of conceptual change is to integrate the piecemeal structure of a conceptual system in a way that increases the internal coherence of the system (diSessa, 1993, 2002, 2004).

Coherence, of course, is as disessa himself writes, “a vague word”, but as he continues, “one important core meaning (of coherence) has inherently to do with relations; that is, the meaning of coherence requires an articulation of structure.” (diSessa, 2008). Even if the term is often left unspecified in the context of conceptual change studies, a useful description for conceptual coherence can be found, for example, from Thagard and Verbeugt (1998, also Thagard et al, 2002). Thagard and Verbeugt defines coherence as follows: (i) Conceptual coherence is a symmetric relation between the pairs of concept, (ii) a concept coheres with another concept if they are positively associated i.e. if there are objects to which they both apply, (iii) the applicability of a concept to an object may be given perceptually or by some other reliable source, (iv) a concept
incoheres with another concept, if they are negatively associated, i.e. if an object falling under one concept tends not to fall under the other concept. Finally (v) the applicability of a concept to an object depends on the applicability of other concepts. Even if Thagard and Verbeugt speak explicitly about coherence of concepts, there is no reason a priori, why their description of coherence could not be applied to the elements of subconceptual systems or more complex entities, such as elements of belief or knowledge systems.

These three approaches are perhaps the most widely accepted descriptions for the task of conceptual change. In an ideal account, these descriptions would be given in an exact way, but at least to our knowledge there are not any exact formulations of conceptual change available. In addition, philosophically speaking, it is still an open question, what is it about fruitfulness, plausibility, utility or coherence that makes the learning task as an instance of conceptual change. Perhaps, very roughly, one might say that conceptual change happens when a student does not merely accumulate more knowledge, but her conceptions of phenomena in a certain domain undergo a restructuring process that affects the conceptual system in a way that increases utility, plausibility or coherence of that system.

**Explanation step 2: The Mechanisms of Conceptual Change.** Now, let’s move to the second step of explanation. This aspect of explanation answers questions like: “how does the mechanism transform the input to generate the output (step by step)?”. In the literature of mechanistic explanations, there are several attempts to specify this notion of “cognitive mechanisms”. For example, Bechtel (2008) defines cognitive mechanisms as follows: A (cognitive) mechanism is a structure performing a (information processing) function in virtue of its components parts, component operations, and their organization. Typically in the case of hard core cognitive explanations, these mechanisms are given descriptions by specifying the precise algorithms or by other formal means.

**The mechanisms of conceptual change.** In the literature, there are many suggestions for the “mechanisms” of conceptual change. For example, Chi talks about categorization and recategorization, while Carey speaks about differentiation, coalescence and bootstrapping. Vosniadou focuses on accommodation and assimilation, Thagard writes about branch jumping and tree jumping, and Ohlsson focuses on resubsumption.

However, often these purported mechanisms of conceptual change are rarely specified with sufficient (computational) detail (for discussion, see Rusansen and Pöyhönen, 2012). The descriptions of these mechanisms are often quite shallow and offer no information about the precise structure of mechanisms. For example, Chi describes conceptual change as a form of recategorization process by saying that “[c]ategorizing is the process of identifying or assigning a concept to category to which it belongs” (Chi 2008, 62), and by writing how “conceptual change is the process of removing misconceptions... (which) are, in fact, misconceptions of concepts” and “conceptual change is merely a process of reassigning or shifting a misconceived concept from one category to another” (Chi, 2008, 62, italics added).

However, Chi offers no description of how “identifying” or “assigning” actually happens, or what kind of cognitive mechanisms they actually are. From an explanatory point of view, this is problematic. Genuinely explanatory models are models, in which the phenomenon is explained by giving an accurate and sufficient description of how a (causal) mechanism, a hierarchical system composed of component parts and their properties sustains or produces the phenomenon (Bechtel and Richardson, 1993; Machamer at al., 2000; Craver, 2006, 2007). In addition, genuine explanations offer the ability to say not merely how the system in fact behaves, but to say how it would behave under a variety of circumstances or interventions (Craver 2000, Craver 2007, Woodward 2003).

So, even if even if these “identifying”, “assignings” and “categorizings” (or any other similar “ings”), were constantly referred as mechanisms, they often fail to satisfy the requirements for genuine mechanism descriptions, because the structure of these mechanisms is not specified in a detail. Instead, often the purported “mechanisms” are, or include, more or less filler terms. Filler terms describe only the relationship between the input and the output of the process, but they offer little specific information of how the change was brought about.

If a mechanistic model is incomplete, and it includes filler terms, it should rather be called a “mechanism sketch” than a genuine explanation (Craver, 2006, 2007). Philosophically speaking, having numerous filler terms in an explanation does not only threaten to undermine its explanatory power, but filler terms may also be barriers to scientific progress when they veil failures of understanding (Craver, 2006, 2007). If, for example, the term “assign” is used to stand for a process with largely unknown properties, then we really do not explain what happens, but in the worst case scenario we may also have only an illusion of explanation (Craver, 2006; Rozenblit & Keil, 2002).

**The details of mechanisms.** In addition, when the details of these mechanisms (reorganisation, bootstrapping, resubsumption, category shifts, etc.) is analyzed, they are often just collections of some more basic cognitive mechanisms (such as categorization, mapping, transfer, assimilation, accomodation, analogical reasoning, inductive inference, abduction and so forth), which are ultimately thought to be responsible for the conceptual change.

For example, in Stellan Ohlsson’s (2009) account conceptual change happens, when a person uses analogical transfer to map conceptual system from one domain A to a new domain B, which has been earlier conceptualized by another system. According to Ohlsson’s model, if the new
system is evaluated to be more useful, the target domain is reinterpreted by it.

As Ohlsson says, the resubsumption theory “does not introduce any cognitive processes that are specific to conceptual change” and “no special purpose cognitive mechanism kicks in to produce conceptual change” (Ohlsson, 2009, p. 32). Instead, resubsumption is simply a process, which involves analogical reasoning, transfer, analogy, transfer, different kinds of mapping and interpretation and all these familiar cognitive mechanisms.

From the explanatory point of view, this is not shocking news. It is quite common, as for example Bechtel and Richardson (1993) emphasize, that complex mechanisms are, and often must be, decomposed into simpler (or more basic) submechanisms that are ultimately responsible for the orchestrated functioning of the higher level mechanism (Bechtel & Richardson, 1993, see also Craver, 2007). However, if the submechanisms are finally doing the explanatory work, they should be given a proper description. If they are not described in a detail, then we really have no explanation as to how conceptual change happens.

**Evaluation of relevance.** Given the complexity of cognitive processes in general, and especially the complexity of conceptual change, in practice it is really difficult to distinguish those underlying submechanisms (attentional-, memory-, reasoning-, mapping mechanisms etc.) that are doing the explanatory work from those which are not. As Ohlsson emphasises (Ohlsson, 2009b, p. 70), a theory of conceptual change just cannot be the list of all possible mechanisms underlying conceptual change, but it must also constraint mechanisms in theoretically principled way. In other words, what we need is a theoretically principled way to evaluate the relevance of submechanisms.

This is, of course, a very difficult demand. Philosophically speaking, one possible line might be to argue that the relevance for a certain mechanism - or certain mechanisms - could be evaluated by knowing how the mechanism’s inputs and outputs interact with their context i.e. by knowing its causal (as opposed to say, intentional relations) with the environment. A natural way to continue this argument would be to refer to the manipulationist account i.e. to argue following Woodward (2003) that those mechanisms are relevant, which do not only have impact on how the cognitive system of a learner in fact behaves, but which have impact also on how it would behave under a variety of circumstances or interventions.

However, it seems to be that in the case of conceptual change - and in genuinely cognitive explanations in general - the explanatory relevance must also be described at least partially by referring the task of the conceptual change as well. A theory of conceptual change should be able to tell, why certain mechanisms are required or are appropriate for achieving conceptual change, and why some other aren’t.

The task level description is needed to characterize representational requirements and constraints for the descriptions of appropriate learning mechanisms. If one thinks that the task should define in terms of utility, then one should characterize those mechanisms that are responsible for “utility making”. However, for doing this, the task level – utility, coherence, intelligibility – must be specified first, and then this specification provides justification to relevance claims concerning the specific mix of concrete mechanisms underlying conceptual change.

**Concluding remarks**

Conceptual change is organizing a multiplicity of learning mechanisms to achieve learning that makes the conceptual system “better” by way of creating new (for the learner) concepts. These concepts do not “pop” into existence in a miraculous way but are typically gradually and sometimes painfully crafted by the cognitive mechanisms from existing material.

According to the mechanistic account, explanation requires that the mechanisms responsible for a certain type of conceptual change should be specified in a detail. This can be really challenging in the case of conceptual change. Conceptual change is a really complex cognitive process, and it may involve a hierarchical collection of many different submechanisms. Some of those are better "known" (categorization, inductive reasoning), some of those aren’t (mapping mechanisms). In addition, there are many different forms of conceptual change, and they may involve several different mechanisms.

However, as also Ohlsson emphasizes (Ohlsson, 2009b), a theory of conceptual change cannot just be the list of all possible mechanisms, but it must also make some constraints for the list of explanatory relevant mechanisms. A theory of conceptual change should be able to tell, why certain mechanisms are required or appropriate for conceptual change, and why some other aren’t. For this reason, the task level also matters. The task level identifies the learning episode as conceptual change by identifying the relevant type of difference between initial state (no concept) and outcome (has concept) is an essential part of explanation because it provides not only the characterization for the explanandum of explanation, but it is also needed to evaluate the explanatory relevance of mechanisms.

**Acknowledgments**

The participants of ConChaMo-workshops. This work was funded by Finnish Academy (project number 1133369).
References


