Counterfactual Imagination as a Mental Tool for Innovation

Monika Chylińska
Institute of Philosophy, Catholic University of Lublin, Poland
moni.chylinska@gmail.com

Received 17 April 2017; accepted 26 September 2017; published 21 November 2017.

Abstract

In the article I demonstrate some of the possible ways by which counterfactual imagination can lead people to innovation and the creation of novel and valuable solutions. I start with adopting the broad definition of counterfactuals, by which counterfactual imagination is understood as the ability to imagine alternative states of affairs which can relate to the past, present or future. I explain how counterfactual imagination differs from other sorts of imaginative and creative thoughts, pointing out that counterfactual types of thinking always rely on facts and involve a change in some features of the actual world, leaving other such features unaltered. I also show that the concept of counterfactual imagination can be useful when we aim to describe the very earliest manifestations of imaginative capacities in children, which can be seen in their make-believe games. All the mentioned characteristics of counterfactual imagination are further used to examine how what if and would be sorts of thinking and imagining might influence people’s creative performance. I conclude with the suggestion that—if guided properly—counterfactual imagination could be a truly valuable mental tool for innovation. This demonstration is partly influenced by Ruth Byrne’s multi-faceted analysis of counterfactual imagination, mainly from her book, The Rational Imagination: How People Create Alternatives to Reality.

Keywords: constraints; counterfactual imagination; creativity; innovation; possibilities.

Introduction

More than a century ago, John Greenleaf Whittier (1898) expressed the folk intuition which is still vivid in our present times: “For all sad words of tongue or pen, the saddest are these: It might have been!” (p. 153). It leads one to sadness and misery when
one thinks about the possible and promising events that could have happened or could have almost happened, which can be a more heart-breaking case, but that which did not happen. “It might have been a masterpiece!” could be piteously thought by a painter who accidentally poured paint on his canvas. On the other hand, as I aim to show in this paper, such trains of thought as it might have been or what if are a powerful tool to expand and even enable our human capacity to innovate and create.

Suppose that another painter also thinks that his finished painting might have been a masterpiece (but was careful enough not to spill the paint). He is looking at his almost-masterpiece and asking a series of what if questions, or visualizing some what if images in his mind.1 “What if I (had) used more blue?” “What if I (had) painted more sunflowers?” In this case, it is not hard to predict some further benefits of such could have been imagining, e.g., we can suppose he is learning from his mistakes or he is planning how to paint a more beautiful still life painting next time. Indeed, what if or might have been thoughts have already been shown as thoughts that can yield helpful scripts for future behavior, including creative behavior (Roese, 1994). Nevertheless, there are many more directions in which we can study how such thoughts (hereinafter called counterfactual thoughts) can be linked to innovation and to our general potential to create.2 In this article I seek to demonstrate some of these connections and several practical applications. This demonstration is influenced by Ruth Byrne’s multi-faceted analysis of counterfactual imagination, mainly from her book, The Rational Imagination: How People Create Alternatives to Reality (2005).

**Counterfactual Imagination That Relies on Facts**

Counterfactual imagination is a distinct ability among our other mental capacities with which we consider alternative states of affairs, or, in other words, generate and process counterfactuals (Byrne, 2005). It is said to appear very early in human development at the age of 18 months, when children start engaging in pretend play (Amsel & Smalley, 2000; Harris, 2000; Lillard, 2001; Weisberg, 2015). Moreover, counterfactual imagination is sometimes perceived as a particular evolutionary precedent and as a potential explanation of the unique human cognitive attributes (De Smedt, 2011; Suddendorf, 2013).

---

1 I support the view according to which thinking what if does not need to be expressed propositionally since alternative states of affairs can be simply imagined or visualized. This will be more precisely explained later.

2 I agree with the proposal of Margaret Boden (2004) and Maria Kronfeldner (2009) to treat as creative such cognition that is a basis for giving new and valuable ideas from the perspective of the individual alone, not only for producing what is new and valuable in an objective or communally agreed upon sense.
Although this phenomenon is discussed in the literature as counterfactual thinking (Roese, 1997) or counterfactual reasoning (Weisberg & Gopnik, 2013), I prefer to refer to it here as counterfactual imagination (referring to Byrne [2005]), thus suggesting that we can not only deliberately think or reason about alternative states of affairs, but we can also imagine or visualize them. Moreover, I do not want to treat this ability only as the ability to construct possible scenarios that have to rely on the past, as did Roese (1997) or Kahneman and Varey (1990). Indeed, most analyses have focused on describing counterfactuals as alternatives for past events which are uttered with the formula “What might have been if X had or had not happened?” I argue that this capacity is a more general one, since I define it as an overall capacity for imagining various alternative scenarios. We can construct counterfactuals not only for what has happened, but also for actual and habitual states of affairs, as well as for events anticipated as likely to happen (future hypotheticals). Additionally, we assert the existence of timeless counterfactuals, as is shown in the example “If kangaroos had no tails, they would topple over” (Lewis, 1973; Woodward, 2011), where time is neither key nor decisive in consideration of this counterfactual. Therefore, I define counterfactual imagination as embracing all thoughts about alternatives to the here and now, including, for example, thinking about the future and pretend or fictional worlds (Amsel & Smalley, 2000; Edgington, 2011; Gopnik, 2009; Harris, 2000; Woodward, 2011).

When applying such a broad definition of counterfactual imagination, it is important to explain how it is described differently from other sorts of imaginative and creative thoughts; for instance, from the notion of divergent thinking. An important characteristic of counterfactual imagination highlighted in literature is that it counters the facts, where the facts part is actually highly significant. Namely, we always refer to reality while imagining alternative states of affairs or—to put it differently—while pondering possible worlds. This means that these counterfactual types of thinking involve a change in some features of the actual world, while other such features are left unchanged (Woodward, 2011). Let us go back to the previous example of the painter who thought “What if I had used more blue in my painting?” Actually, there is only one thing that he alters in his visualization: the amount of blue paint. The other features of the actual world (such as the portrayed objects or the size of the canvas) are left unaltered.

3 Such counterfactuals about the past can be called real-world counterfactuals here and there (Beck & Riggs, 2014).

4 My viewpoint on this matter is similar to the standard broad view of counterfactuals in which it is claimed that, as counterfactuals, we should perceive not only sentences expressed in the subjunctive mood, but also selected sentences uttered in the indicative mood, such as “If I finish this paper, I will go for ice cream” (see Hoerl, McCormack, & Beck, 2011; Woodward, 2011).
Furthermore, counterfactual imagination can be perceived as ever-constrained by some facts or commons, or by particular external and internal conditions (Byrne, 2005).\textsuperscript{5} It may be constrained by the need to create minimal changes to reality (Byrne, 1996), or that reality must be recoverable from the imagined alternative (Byrne, 1997). Otherwise, our imagined worlds would be incomprehensible, senseless, or perhaps even not possible to imagine at all because people may understand a counterfactual alternative only through the lens of the facts from which it was created (Byrne, 2005, 2016).

**Counterfactual Imagination in Human Development**

It is not news to say that children can be highly imaginative; this is mostly because from the age of 2 they engage in make-believe games (Fein, 1981). Talking especially about counterfactual imagination, we could find a number of studies in which pretend play is described as an activity based on counterfactual-like cognitive capacities such as (a) the competence in pondering possibilities (e.g., through a mental possible worlds box mechanism [Nichols & Stich, 2000]), (b) suppositional thinking (Perner, 1991), or (c) the mechanism of decoupling, thus disconnecting primary representations from their real functions (Leslie, 1987, 1994). According to Leslie (1987), a psychologist broadly recognized for his research on pretend play, in make-believe contexts children create a reality that is an alternative to the one known by them or believed to be true. In this sense, humans at the very beginning of their development are capable of cognitive innovation—with their peculiar imaginative competence they can transcend time, place, and/or circumstance to think about what might have been, design the future, create fictional worlds, and consider alternatives to the actual experiences of their lives (Harris, 2000).

Some researchers have recognized the counterfactual features of children’s pretense episodes in a more explicit way (Amsel & Smalley, 2000; Harris, 2000; Weisberg, 2015; Weisberg & Gopnik, 2013). For instance, they say that pretending—as with counterfactual thinking—involves representing possible worlds (Weisberg, 2016). According to Angeline Lillard (2001), performing make-believe games serves a similar function for children, in some respects, as the Twin Earth construct serves for philosophers: it enables them to imagine and reason about an alternative world which resembles reality in some parameters. In other words, while pretending a

---

\textsuperscript{5} Here we see how the notion of counterfactual imagination is contrary to the famous view of imagination proposed by Kant (1781/2000), who claimed that our imagination is completely free from reality: “For the imagination (as a productive cognitive power) is very mighty when it creates, as it were, another nature out of the material that actual nature gives it . . . In this process we feel our freedom from the law of association . . . for although it is under that law that nature lends us material, yet we can process that material into something quite different, namely, into something that surpasses nature” (p. 192).
child is considering and acting out certain alternative scenarios in which some features of reality change, and some do not. Therefore, *ipso facto*, during their play children experience some imaginative constraints and follow some thinking rules.

Although our first impulse could be to claim that children’s pretend play (as with imagination) is unlimited and freely detached from reality; however, it is subject to certain regularities. A child does not pretend in a completely free-flowing manner, but preferably by following selected rules and by making his pretense understandable for others⁶ (Bretherton, 1989). A child creates alternatives to existing events which, nevertheless, retain causal powers of mundane reality (Harris, 2000). Thanks to this pretense, an imaginative activity in which children posit the existence of fictitious worlds could paradoxically be an important tool for acquiring knowledge about the actual world⁷ (Weisberg & Gopnik, 2013).

**Applications of Counterfactual Imagination**

As my previous analysis shows, one aspect of counterfactual imagination is that it is constrained by facts. Would it be disturbing to say that counterfactual imagination is not as free as we first thought and that it does not help us to act as highly innovative and creative humans?

From some studies and applications of counterfactual imagination in the creative process, we learn that this specific ability of thinking *what if* or *what might have been* can enhance our creativity and innovation under certain conditions. For example, Markman, Lindberg, Kray, & Galinsky (2007) showed that so-called additive counterfactuals, which add new antecedent elements to reconstruct reality (e.g., “If only I owned an umbrella, I would not have gotten wet”), can evoke an expansive processing style. This processing facilitates creative generation or divergent thinking. On the other hand, subtractive counterfactuals (e.g., “If only it hadn’t rained today, I would not have gotten wet”) evoke a relational processing style that simplifies analytic task performance. Hence, additive counterfactuals can be viewed as more creative—they can “go beyond the original premise set, fabricating novel options perhaps never considered in the past” (Roese, 1994, p. 807). Creative leaders can then encourage people to structure their counterfactual stories to be more additive in order to allow them to be open to describing alternative states of affairs which were not part of the factual event. By these means we might boost the probability of creative and innovative performances in individuals.

---

⁶ If interested in the topic, see an illuminating short movie on children’s pretense: https://aeon.co/videos/children-at-play-provide-a-rare-glimpse-into-the-imagination-ours-and-theirs

⁷ There have been a number of studies on the function that pretend play may serve for children’s development, among which the most prominent are the studies on the cognitive connections between pretending and understanding the minds of the others (Gut & Wilczewski, 2016; Leslie, 1994; Lillard, 2001).
Considering the previous applications of counterfactual imagination in innovation as well as in various creativity training programs which do not have to lead to any specific innovative products, we see that there is one well-known method that uses counterfactuals explicitly. Among lots of innovative strategies, there is one based on asking several what if questions and searching for original would be answers. “What if another advanced species existed?”, “What if there were no gravity?”, “What if we saw only one color: red?” Such fictitious situations prompt our creative thinking by letting us imagine different possible worlds in which some features do not have to be the same as in the actual world. Hence, this counterfactual method introduces us to the mental state of creative generation of novel, non-actual ideas. At the same time, asking what if can be fruitful and is often adopted at the exploratory stage of creative and innovative processes (according to Geneplore model of creativity by Ward, Smith, & Finke [1999]). When we see a possible solution for a problem or when we construct a prototype version for our innovative product, we might ask a series of what if questions (e.g., “What if I made it a bit smaller?”) to explore more possibilities to find the most creative one or to ensure that our final proposal is appropriate, or as valuable as it could be. Therefore, imagining what would be is important both at the preliminary generative stage of innovation and at the exploratory stage, when we evaluate and improve the selected ideas or artefacts. Counterfactual imaginative strategies are used in innovation in many ways and during various periods of the creative process.

**Counterfactual Imagination as a Mental Tool for Innovation**

To complement the above brief notes on the adoption of the notion of counterfactuals in the creative field, in this section I will follow and develop the analysis of Byrne (2005), who stated that counterfactual imagination shares many similarities with other sorts of creative thought.

Byrne’s main idea is that counterfactual imagination should be viewed as rational, since people create counterfactual alternatives by thinking about possibilities guided by a small set of principles. For example, when people create a counterfactual alternative to reality by mentally altering some aspects of reality, some of these aspects, which Byrne calls ‘the fault lines of reality,’ seem more mutable than others. That is, they are readily changed in a mental simulation of an event (Kahneman & Miller, 1986). Time seems mutable in our imaginary scenarios, while our own existence seems not. For example, we would say “If I had no time, I could not write this article” rather than “If I did

---

8 Notice that there are at least two possible ways of answering such what if questions: (1) as many different answers as possible (known as divergent thinking) or (2) to explore and develop one answer, e.g., What if kangaroos had no tails? They would topple over and then they could not move so fast (so they would be easy victims for predators, which means they could be in danger of extinction etc.). I would call this second way of processing cascading thinking and would relate it at the same time to exploratory and generative ways of creative problem solving (referring to the Geneplore model of creativity by Ward, Smith, & Finke [1999]).
not exist, I could not write this article." Different people tend to change the same features of the actual world when they think about counterfactual events; therefore, there are remarkable similarities in what everyone imagines. Moreover, thinking about other principles that lead our imagination, Byrne points out that we usually think of just a few (if not only one or two) possibilities to mentally represent some aspects of reality. We also do not tend to focus on impossibilities, that is, things that could never have happened given the way the world is, e.g., “If aliens attacked the city, I could not write this article.” Likewise, we do not explain things by mentally altering natural laws; we do not say “It would be harder to write this article if there was too much oxygen in the air” or “She would not have fallen from the ladder if there had not been gravity” (Seelaü, Seelau, Wells, & Windschitl, 1995).

All these described principles of the counterfactual imagination seem not to be too friendly to creativity and innovation. Claiming that this type of imaginative thinking shares similar mental features to creative thinking would result in saying that different people guided by similar imaginative rules tend to create similar things when they are situated in the same context to solve. Nevertheless, this is exactly the case. For instance, when different people have been asked to draw an alien from another planet, they tended to picture their cosmic creatures with sensory organs, most commonly eyes, and also with the limbs such as legs (Ward, 1994). People may think of just a single possibility to mentally represent some aspects of reality, such as the presence of eyes or the law of gravity, both of which seem immutable in the sense they are often left unaltered in our imagination.

Does this all mean that counterfactual imagination is counter to creativity and innovation? Not at all. On the contrary, counterfactual imagination can lead us to highly original and valuable ideas and objects. Accordingly, I see some possible paths of interpreting Byrne’s studies in favor of the creative power of the counterfactual imagination.

The main path I will sketch here could be called the meta-cognitive one. It shows that it is helpful for one to have and develop original and novel ideas when one has a sort of meta-knowledge of counterfactual imagination. (Of course, Byrne’s principles are taken into account.) To be innovative, we could try to break some of the imaginative rules, as most creative nonconformists usually do. These are the possible ways to do it:

1. Instead of making minimal changes to reality in your imagination, think bigger and alter more.
2. Do not rely solely on the elements of your conceptual knowledge that first come to your mind when creating: add more alternatives, take a look around, travel to a new place, etc.
3. Instead of using ‘close counterfactuals’ which are close to reality, (“If I finish the paper, I will go to sleep”), try to use remote ones (“If I finish the paper, I will go to visit one of the possible worlds”).
4. Make mutable what seems to be immutable, e.g., imagine an alien without eyes and legs (!).
5. Try to change the diagnostic aspects of a concept which seem least mutable, e.g., if prickly spines are typical of a cactus, think about a cactus that does not have them, or imagine shoes without soles.
6. Alter the laws of nature when trying to be innovative.9

Following these suggestions, we could make use of counterfactual imagination at its best in our innovative and creative performances. To put it differently, good imagination management could lead us to novel and original solutions, and this would be the evidence that counterfactual imagination is a truly valuable mental tool for innovation.

One Last Thing of Importance

Finally, I will point out the most basic and fundamental function that counterfactual imagination might serve for our creativity and innovation. Namely, it facilitates us to alter reality and to invent anything with a touch of or a basis in the actual world, for all inventions consist of alternations or developments of reality. Why in the Upper Paleolithic era, around 10,000 years ago, did an extraordinary cognitive revolution in human culture occur? According to Harris (2000), this was caused by the appearance in humans of the capacity to invoke imaginary possibilities. Such a cognitive revolution of counterfactual imagination enabled the invention of tools and new strategies for hunting. In contemporary times it enables us every day to think up new ideas and to invent either simple or complicated useful objects such as teapots or robots. To be creative, we certainly need to guide and use our counterfactual imagination properly, as how revealing it is can be seen from Erin Hanson’s question:

“What if I fall? Oh darling, what if you fly?”

Acknowledgments

Sincere thanks to my friend Claire Zhenxu Fan as well as to Hannah Drayson, Raluca Briazu, Abigail Jackson, Klara Łucznik and the other researchers at CogNovo whose comments and suggestions helped me to develop and improve this paper.

Work on this paper was supported by the National Science Centre, Poland (UMO-2016/21/N/HS1/03495; title of the research project: Counterfactual Imagination and Pretend Play: The Cognitive Underpinnings of Human Creativity).

9 There is a truly original children’s book about a boy who defies the laws of gravity: The Terrible Thing That Happened to Barnaby Brocket by John Boyne.
References


Response to “Counterfactual Imagination as a Mental Tool for Innovation”  
by Hannah Drayson

The paper proposes a number of ways to use counterfactual imagination as a method in design practice. In the following, I refer to a number of complementary references for this area of enquiry from literary theory and science and technology studies.

First, the observation that counterfactual thinking is conformist suggests a connection to writing in literary theory on the topic of alternative worlds. This work considers the way in which the familiar and unfamiliar are held together in the practice of creating complex alternative scenarios. As science fiction scholar Darko Suvin defines it, sf is “a literary genre . . . whose main formal device is an imaginative framework, alternative to the author’s environment” (Suvin, 1979, pp. 7–8). Some of the concepts generated to analyze these writing practices might offer another dimension to the discussion, for example: Suvin’s theory of the sf novum, a single, scientifically plausible innovation around which a story is told and which produces an experience of “cognitive estrangement” (inviting speculation in the reader about possible alternatives to the status quo). Other work in this vein connects with thinking about the relation of the counterfactual to creative imagination; according to Arne Zettersten (2011), J. R. Tolkien’s belief in myths was as transcendent truths (complicating the definition of factual and counterfactual).

This paper appears to raise a definition of counterfactual imagination that invites further questions, possibly because of the slippery meaning of the word “fact.” If—as it is defined here (p. 243)—“counterfactual imagination [embraces] all thought about alternatives to the here and now,” then where does factual information that is not about the “here and now” fit into this definition? The difficult processes by which matters of fact come to acceptance are an object of study, as much in the humanities (Latour & Woolgar, 1979) as in the sciences. However, this definition seems to assume that what we directly experience and what is “true” are the same thing. Take, as an example, a “fact” that kangaroos have long tails. In terms of my experience of the “here and now,” there is very little to distinguish this “fact” from other thoughts or imaginings that I might entertain, and what of thoughts that I have which are mistaken? With the paper’s goal to instrumentalize counterfactuals for the purposes of enhancing creative innovation—the definition of a counterfactual thought may have to include the coda that it is counter to the facts as understood by the individual who is imagining them. This individual perspective may offer a further line of exploration.
References

