Logic and **L**ogical **P**hilosophy
Volume 14 (2005), 69–88
DOI: 10.12775/LLP.2005.005

Tim De Mey*

TALES OF THE UNEXPECTED

Incongruity-Resolution in Humor Comprehension, Scientific Discovery and Thought Experimentation

Abstract. Some scholars suspect that thought experiments have something in common with jokes. Moreover, Thomas Kuhn has suggested that what happens to someone who thinks through a thought experiment "is very similar to what happens to a man, like Lavoisier, who must assimilate the result of a new unexpected experimental discovery" (1964: 321). In this paper, I pinpoint the presumed commonalities. I identify, more specifically, what cognitive linguists call "incongruity-resolution" as the problem-solving process not only involved in humor comprehension, but in scientific discovery and thought experimentation as well.

1. Introduction

What do jokes, scientific discoveries and thought experiments have in common? This sounds as the beginning of a joke and—given that the objects

*The author is a Postdoctoral Fellow of the Fund for Scientific Research – Flanders (Belgium) and works at the Centre for Logic and Philosophy of Science – Ghent University. Thanks to Tony Veale (UCD), Geert Brône and Kurt Feyaerts (KUL), Diderik Batens, Kristof De Clercq, Joke Meheus and Liza Verhoeven (UG) for informal discussions on the subject and to Erik Weber and Dagmar Provijn for comments on the draft version of this paper. The research for this paper benefited from a bilateral scientific exchange project funded by the Ministry of the Flemish Community (project BIL01/80) and the Polish State Committee for Scientific Research.

Received October 31, 2004
 ISSN: 1425-3305

involved are, at least at face value, so disparate—perhaps even a promising one. However, I should warn the reader from the outset that the answer developed and defended in this paper is not really meant to trigger laughter. Framing the basic question in terms of "what do x, y and z have in common?" merely sets the pace in that one typically does not expect the answer to reveal some *essential ingredient* of all things called x, y or z (let alone a sufficient ingredient).[1] What I'm after in this paper is rather a feature that is shared by *many* jokes on the one hand and *some* salient examples of scientific discoveries and thought experiments on the other.

Given this disclaimer, the idea developed and defended in this paper is that it is the *reasoning processes* involved in humor comprehension, scientific discovery and thought experimentation that are quite similar. Some scholars have already hinted at such commonalities. Koestler (1964) famously suggested that humor and scientific discovery have something in common. Kuhn wrote that what happens to someone who thinks through a thought experiment "is very similar to what happens to a man, like Lavoisier, who must assimilate the result of a new unexpected experimental discovery" (1964: 321). And, more recently, Buckley (2003: 22–23) mentioned in passing thought experiments and, more specifically, textbook examples of *reductio ad absurdum*, while discussing humor comprehension.

Unfortunately, however, these suggestions are rather vague. The purpose of this paper is to bring the loose ends together and to identify the problem-solving process at issue as what cognitive linguists call "incongruity-resolution". An incongruity is, very roughly, the cognitive conflict that arises when something *unexpected* happens or is being said.[2] Now, typically, expectations come in different degrees of explicitness, but even when they are

[1] As a matter of fact, such "what do x and y have in common?"—jokes often work by virtue of answering the question by putting a completely superfluous commonality to the fore, thereby making fun, say on a meta-level, of our desire to look for essential features.

[2] A classical *locus* is the paper by Jerome Bruner and Leo Postman on the perception of incongruity in which the problem of incongruity is defined in terms of "the perceptual events which occur when perceptual expectancies fail of confirmation" (1949: 208). On the basis of a "trick card"-experiment, Bruner and Postman were able to show that "for as long as possible and by whatever means available, the organism will ward off the perception of the unexpected, those things which do not fit his or her prevailing set" (1949: 208). Interestingly, in his *The Structure of Scientific Revolutions*, Thomas Kuhn (1962/1970) cited the Bruner-Postman experiment to argue that scientists are often subject to what Bruner and Postman call "dominance", i.e. the denial of incongruous elements. According to Kuhn, scientists typically fail to perceive *unexpected* anomalies, in the sense of anomalies which have not been predicted by the prevailing theory.

outspoken, it will require a considerable amount of what computational linguists call "backward chaining",[3] to pinpoint in what previously happened or was being said, those elements that made one expect something else altogether.[4] In turn, however, this backward chaining *significantly facilitates* the subsequent "forward chaining", i.e. the reasoning towards a solution of the problem in the form of an interpretation, a model or a theory on the basis of which one *could* have expected after all whatever happened or was being said at the time the incongruity initially arose.[5]

Obviously, in the absence of a satisfactory explication of "incongruity" and its "resolution", talk of "incongruity-resolution" remains almost as vague as the analogies developed by Koestler, Kuhn and Buckley. Nevertheless, documenting the variety of domains in which something unexpected triggers a truly *creative* reasoning process,[6] can significantly add up to the goal of getting a firmer grip on incongruity-resolution, for it will allow us to integrate findings from disciplines as diverse as cognitive and computational linguistics at one end of the spectrum and the philosophy and psychology of science at the other.

I will develop my argument as follows. In Section 2 I will comment on Graham Ritchie's (1999) state of the art paper concerning incongruity-

[3] I deliberately imply that "backward chaining" is *a fortiori* involved in the process of explicating the expectations which are violated at the time the incongruity arises. However, I do not believe that incongruity-resolution *necessarily* involves explicating the violated expectations. What it does require, at the very minimum, is that the *source* of the expectations is pinpointed, so those elements that made one expect something else than what one got at the time the incongruity arose.

[4] Here I already seem to subscribe to the "forced reinterpretation"-model of incongruity-resolution, but in Section 2 I will also discuss and integrate findings from the alternative to it, i.e. Suls' "two-stage model".

[5] This outruns the traditional conceptions of "incongruity-resolution" slightly. It is, however, what I have in mind and what I will sketch out and apply in the rest of this paper.

[6] Although I will not have much to say on creativity in this paper, it is obviously a binding factor. I deploy the term roughly in accordance with Herbert Simon's procedural account of it, which evades the conceptual caveats associated with identifying instances of creativity purely or primarily on the basis of the properties of the products (Simon 1977). At the very least, Simon does not require the product to be "truly new" or, say, "objectively valuable". For him, the programs described by Langley and others (1987), which can rediscover Kepler's and Ohm's laws, are as creative as Kepler's and Ohm's original problem-solving processes. Novelty does play an important role in Simon's conception of creativity. However, the novelty he is after is not that of the product, but rather that of the ideas which are shaped during the course of a problem-solving process. Now, as this paper will testify, incongruity-resolution is clearly "creative" in Simon's sense.

resolution theories of humor comprehension. On that basis, I will be able to advance a *tentative explication* of both "incongruity" and its "resolution" that integrates what I take to be the most valuable insights from the two opposing models, i.e. the prevalent forced-reinterpretation model and Suls' two-stage model. Subsequently, I will use that explication of "incongruity-resolution" to describe in Section 3 how William Harvey discovered blood circulation (at least according to his own account of the discovery process in *De motu cordis*, i.e. the famous quantitative argument). I will also imply that similar problem-solving processes are involved in other cases of scientific discovery, the only requirement being that the actual discovery process is touched off by some unexpected observation. In Section 4 I will take incongruity-resolution one step further and apply it to the reasoning involved in thinking through the kind of thought experiments which Thomas Kuhn discusses in his (1964) paper "A Function for Thought Experiments", i.e. thought experiments that trigger conceptual change. As a bonus, I will speculate in the concluding section about the lines along which the explication of incongruity-resolution tentatively advanced in this paper might be further developed.

2. Incongruity-resolution in humor comprehension

Showing that incongruity-resolution is involved in humor comprehension is definitely the most easy part of this paper, for the very idea of incongruity-resolution has been developed by philosophers, linguists and psychologists to capture what it takes, cognitively speaking, to get a joke.[7] Relative to the fact that most contemporary scholars agree that humor relies on incongruity, it is surprising that they haven't yet agreed on a sufficiently precise definition of it. As Ritchie (1999) observes in his state of the art paper:

> The main problem, if we are to develop a detailed theory of verbally expressed humour, is that the notion of 'incongruity' is not clearly defined, and it is not even obvious that all the writers on this subject have exactly the same concept in mind. Ritchie 1999: 78

[7] The roots of these contemporary theories can be found in early-modern philosophical books, notably in Kant's *Critique of Judgment* and Schopenhauer's *The World as Will and Idea*. Kant's and Schopenhauer's definitions of laughter constituted a serious alternative to the predominant superiority theory of humor, which can be traced back to Plato and Aristotle, but was most forcefully put to the fore by Hobbes. See Buckley (2003) for a contemporary defense of the superiority theory. See Morreall (1987) for a selection of texts about humor from the philosophical literature.

Ritchie then goes on to document that the writers on the subject *do* conceive of "incongruity" quite differently. Nevertheless, there seems to be some common ground. Incongruity-resolution theorists typically (a) differentiate between the set-up and the punch line of a joke, (b) claim that the punch line does not make immediate sense to the cognitive agent, and (c) assume that, subsequently, the cognitive agent somehow finds a "resolution" which allows the punchline to be congruous.[8] Take, e.g., the following joke, analyzed *ad nausseam* in the relevant literature:

> (1) "Is the doctor at home?" the patient asked in his bronchial whisper. "No", the doctor's young and pretty wife whispered in reply. "Come right in".

So, according to incongruity-resolution theorists, at least those subscribing to the prevalent forced reinterpretation account of incongruity-resolution (a) everything preceding the punch line "come right in" constitutes the set-up of the joke or, in Godkewitsch's terminology, the joke body, (b) the punch line is somehow at pars with the cognitive agent's interpretation of the joke body and (c) that "incongruity" triggers a problem-solving process in which an alternative interpretation, i.e. one which minimally accommodates the meaning of the punch line, is constructed. So the basic idea of the forced reinterpretation account is that the meaning of the punch line does not only *define* the problem, i.e. the incongruity, but that it also triggers its solution by *evoking* the alternative, less obvious interpretation. Raskin's Semantic Script Theory of Humor adds the notion of "scripts" to this analysis. In example (1) then, the joke body activates the DOCTOR-PATIENT script, but the punch line forces the cognitive agent to backtrack and reinterpret the text in a LOVER script.

In a recent paper, Veale (2004) questions whether the incongruity really *causes* backtracking and reinterpretation. Considering (1), Veale would note that it is possible to give the text a non-humorous rendering, e.g. by interpreting the punch line as an invitation by the doctor's wife to come into the house and await the doctor in the waiting room. So the humorousness of (1) is not as much a matter of incongruity-resolution as, e.g., Raskin would

[8] Despite earlier attempts to mark the difference (Raskin 1985; Attardo and Raskin 1991), the Semantic Script-based Theory of Humor is nowadays commonly classified as an incongruity-resolution theory. Recently, Brône and Feyaerts (2003) have argued that Giora's "marked informativeness requirement" and "optimal innovation hypothesis" fit into the category of incongruity-resolution theories of humor as well. See Attardo (1994) and Ritchie (2004) for comprehensive overviews of linguistic theories of humor.

hold. Rather, the cognitive agent "opportunistically constructs" a humorous interpretation of it. The cognitive agent is not "forced" to backtrack and reinterpret, he or she eagerly "decides" to construct a humorous interpretation.[9]

Suls' two-stage-account of incongruity-resolution constitutes an alternative to the forced reinterpretation-account that can evade the difficulties associated with postulating backtracking (1972; 1983). Ritchie summarizes the basic idea of Suls' account as follows:

> The punchline creates incongruity, and then a *cognitive rule* must be found which enables the content of the punchline to follow naturally from the information established in the set-up. Ricthie 2004: 59

Now, according to Ritchie, the main difference between the two accounts is that they cover different subclasses of jokes: on the forced reinterpretation-model the set-up needs to be ambiguous, but Suls' two-stage model doesn't require anything like that. Ritchie also claims that whereas the forced reinterpretation-model decomposes the humorous effect into slightly simpler concepts, Suls' model "relies on some (undefined) form of 'humorous logic', and so leaves the difficult problem of 'incongruity' relatively untouched" (Ricthie 1999: 83).

However, this evaluation is not entirely fair. Suls has an outspoken conception of "incongruity" and even one that has the virtue of being at once intuitive (to a certain extent) and fruitful (in that it allows for formalization). It is intuitive because it starts from the mundane conception of incongruities as "violated expectations". Here is Suls' definition: "Incongruity of the joke's ending refers to how much the punch line violates the recipient's expectations" (Suls 1972: 92). Now, as Ritchie also indicates, translating one vague notion (i.e., incongruity) into a notion which also lacks a precise definition (i.e., expectation) is not much of a gain. There is, however, one way to get a

[9] Veale's objections are undoubtedly challenging and they pinpoint a weak spot in forced reinterpretation-accounts of incongruity-resolution. However, a sharp distinction between speaker and listener or, perhaps less confusingly, between humor production and humor comprehension can take the spell out of them. Salient examples of what Veale calls "opportunistically constructing" a humorous interpretation, like trumping, i.e. deliberately taking a figurative meaning literally, will more readily be identified as examples of humour production. But one can also consider the choice which the cognitive agent makes, when he or she rules out non-humorous renderings of (1), as his or her contribution to the production of humour. What the forced reinterpretation-account of incongruity-resolution tries to capture, by contrast, is what goes on in the mind of someone who comprehends humour, and not in that of someone who produces humour.

very firm grip on expectations and that it is in terms of "predictions". And that is exactly how Suls conceives of humorous incongruities: what happens at the punch line is not what the cognitive agent predicted would happen. Here is an example of a joke (taken from both Suls and Ritchie):

> (2) O'Riley was on trial for armed robbery. The jury came out and announced, "Not guilty." "Wonderful," said O'Riley, "does that mean I can keep the money?"

So, according to Suls, the cognitive agent predicts, say just before the punch line, that O'Riley will say "does that mean I can go now?" and that outspoken expectation is subsequently violated by the punch line "does that mean I can keep the money" . I agree with Ritchie that "the argument that there is CONFLICT with some specific PREDICTION, and that this causes the humorous effect, is unsupported" (1999: 81). It is even counterintuitive: our expectations do not seem that precise, at least not in this case.

My suggestion for a more formal model of humor comprehension would be to define "incongruity" in terms of "violated expectations", but to resist, *pace* Suls, the temptation of then defining "expectations" in terms of specific "predictions". For our purposes, "predictability" suffices: all we need to account for *incongruity recognition* is that the punch line was not predictable, *could not be predicted* or *derived* on the basis of the first interpretation or rendering. Thus defined, we can also get a more formal grip on what the "resolution" of the incongruity amounts to: confronted with the punch line, the task of the cognitive agent is "simply" to find an alternative interpretation or rendering of the set-up on the basis of which, retrospectively, the punch line *could have been predicted* (if only he or she had not been lead astray during set-up).

3. Incongruity-resolution in scientific discovery

Although explications of the reasoning processes involved in getting incongruity—based jokes are of course important in their own right, they can also be helpful to obtain, relative to the existing models, a firmer grip on the more complex problem-solving processes involved in scientific discovery.[10]

[10] A classical *locus* is the introduction to and the papers collected in Nickles (1980). Giving an overview of the existing models clearly falls outside the scope of this paper. Nevertheless, it might be useful to spell out that as far as the choice is between, on the one hand, the view that discovery occurs in an instant, and on the other hand, the

As I take it, cases in which a discovery process is triggered by some unexpected observation, resemble the incongruity-resolution involved in humor comprehension most. Thomas Kuhn (1962) has already analyzed Lavoisier's discovery of oxygen along those lines, though Kuhn obviously doesn't use the "incongruity-resolution" or "forced reinterpretation" terminology. Here, I would like to analyze a case quite similar to Lavoisier's, i.e. Harvey's discovery of blood circulation, explicitly in terms of incongruity-resolution.

Although the significance of the quantitative argument in the actual discovery process remains a matter of controversy among Harvey scholars (cf. infra), none of them doubts its importance in the reception process.[11] On the basis of the quantitative argument, Harvey was able to convince his contemporaries in a period of *merely* thirty years, to put Galen's long-standing theory aside and to embrace the completely novel idea of blood circulation.[12]

To appreciate the effect of Harvey's argument, it is useful to summarize the theory it challenges. According to Galen, the venal system and the arterial system are almost completely separate. The venal system, which has the liver as its main organ, contains and distributes nourished blood. The arterial system, which has the heart as its main organ, contains and distributes vivified blood. Now the various parts of our body require both nourished and vivified blood to function properly. That inevitably requires that there is in both the veins and the arteries (a steady) blood flow in both directions (i.e., towards and away from the heart). Another contention of Galen's model is that both kinds of blood are actually consumed by the various parts of the body. So the question then is how these kinds of blood are produced in sufficient amounts. According to Galen, that is the task of the liver: it turns *chyle* (a derivate of digested food) into blood.

Interestingly, Harvey's argument takes the form of a report on how his research proceeded (which perhaps explains at once why some Harvey schol-

view that it involves a process during which *local problems* are solved, i.e. anomalies and contradictions that call for resolution by way of reformulation, I side with the latter.

[11] The importance of the quantitative argument in the reception process is striking, since both opponents and proponents had a hard time figuring out what kind of argument it was: "it had no place in the contemporary categorization of the modes of reaching knowledge" (French 1994: 93). Elsewhere, I have suggested that the quantitative argument is a thought experimental-argument (De Mey 2003), but I will not press the point here, since the intuition underlying that classificatory judgment will become clear when I analyze, in section 4, thought experimentation in terms of incongruity-resolution.

[12] Holmes (2002: 173), at least, argues that both Harvey's discovery of blood circulation and Lavoisier's discovery of oxygen were *comparatively quickly* accepted.

ars continue to be suspicious about its actual role, for scientists tend to idealize the way in which they make their discoveries). Here is the famous opening paragraph of Chapter 8 of *De motu cordis*:

> Now truly, when I had many times and seriously considered with myself the varied means of searching, and how many there were! both from the dissection of living creatures for experiment's sake, and from the opening of arteries, as well as from the symmetry and great size of the ventricles of the heart and of the vessels which go into it and go out from it (for Nature who makes nothing in vain would not have allotted to those vessels so comparatively large a size to no purpose), and from the carefully balanced and exquisite contrivance of the valves and fibres and from the rest of the fabric of the heart, and from many other things, and when I had for a great while turned over in my mind these questions, namely, how great was the abundance of the blood that was passed through and in how short a time that transmission was done, and when I had perceived that the juice of the food that had been eaten could not suffice to supply the amount of the blood—nay, more, we would have veins empty and altogether drained dry and arteries, on the other hand, burst open with the too great inthrusting of blood, unless this blood should somehow flow back out of the arteries once more into the veins and return to the right ventricle of the heart—I began to bethink myself whether it might not have a kind of movement as it were in a circle. And this I afterwards found to be true, [...].
>
> Harvey 1578/1976: 74–75

So Harvey seems to explain how surprised he was initially to find that, during contraction, the heart forcefully ejects such a large amount of blood into the aorta and that it took him quite a long time and elaborate further research to pinpoint the problem, i.e. to isolate, within Galen's sophisticated model, the elements that made him expect a much smaller amount of blood: "when I had for a great while turned over in my mind these questions, namely, how great was the abundance of the blood that was passed through and in how short a time that transmission was done [...]".

After having pinpointed, however, that the precise basis of his expectancy was Galen's idea that blood is continuously produced by the liver on the basis of nourishment and continuously consumed by the rest of the body, he was able to turn his initial surprise into a straightforward contradiction: "[...] and when I had perceived that the juice of the food that had been eaten could not suffice to supply the amount of the blood-nay, more, we would have veins empty and altogether drained dry and arteries, on the other hand, burst open with too great inthrusting of blood [...]". So if

both Galen's theory and his own estimations about the amount of blood ejected into the aorta during contraction were right, the veins would soon be emptied and the arteries would simply burst with too much intrusion of blood.

Part of the rhetorical force of Harvey's argument is of course that even on much smaller amounts, that is exactly what would occur.[13] For my purposes, however, it is more important to note that after having explicitly formulated the contradiction, Harvey immediately adds that blood circulation is the only way-out, i.e. the only explanation of why, given such or such an amount of blood, the veins are not emptied and the arteries do not burst:[14] "unless this blood should somehow flow back out of the arteries once more into the veins and return to the right ventricle of the heart [...]".

I already mentioned above that the status of the quantitative argument is a controversial subject among Harvey scholars. Opinions range from the view that Harvey's quantitative considerations really triggered the very idea of blood circulation to the view that the quantitative argument is purely *post hoc*, carefully designed for the sole purpose of convincing his contemporaries. The historiographical problem arises from the fact that Harvey allegedly reported to Boyle that the discovery was touched off by his reconsideration of the function of the venous valves, which only allow for blood flow in the direction of the heart. On top of that, Harvey also performed experiments with ligatures, showing unidirectional blood flow in both veins and arteries, the former towards and the latter away from the heart. Although Harvey deployed these three elements to argue for blood circulation in *De motu cordis*, Harvey scholars still disagree on their respective weight in the discovery process and especially on the order in which they played that role. Bylebyl (1982) argues, e.g., that the direction of the venous valves

[13] As a matter of fact, shortly after the paragraph quoted above, Harvey goes on with stressing the scope of his argument, by inviting the reader to change to suppose much smaller amounts. E.g., varying the number of pulsations, he writes: "But, grant that it be not done in one half hour, but in one hour or in one day, whatever you choose, it is clearly evident that more blood is incessantly driven through the heart by its pulsation than can possibly be either supplied by the food we eat or contained at one time in the veins" (Harvey 1578/1976: 80).

[14] So the contradiction has a dual function. Firstly, it serves to "destroy" the old theory. In that sense, Harvey's line of reasoning can be considered as a classical *reductio ad absurdum*: from the old theory and a novel observation (i.e., the amount) a contradiction is derived. But the contradiction, or rather the backward chaining leading up to it, also puts restrictions on the possible new theories. So on top of its destructive function, it has a constructive function.

hardly fulfilled any function in the discovery process. Whitteridge (1971), by contrast, attributes more significance to the venous valves and even suggests that it came first, i.e. before Harvey's quantitative considerations. According to Pagel (1976), on the other hand, Harvey proceeded just the other way around.

I tend to side with those who claim that Harvey's quantitative considerations preceded both the reconsideration of the function of the venous valves and the experiments with ligatures. After all, as French (1994) explains, Harvey first innovation to the field was his thesis that the active moment in a heartbeat is during contraction (or systole) and not, as Galen suggested, during expansion (or diastole).[15] This thesis met a lot of criticism which forced Harvey in turn to look in more detail to the force with which and the amount in which the heart propels blood into the aorta. Given that Harvey formulated this thesis almost at the very beginning of his investigations and that the criticism followed immediately, it makes sense to presuppose that Harvey was already surprised or even shocked to find such a large amount of blood forcefully propelled in the aorta at an early stage in his investigations:

> Let us recall that Harvey was seeking to defend his thesis of the forceful systole by showing that blood emerged from the heart with some force and in some quantity. [...] Clearly in a comparatively short space of time the whole of the blood contained in the body must pass through the heart. When this realization first came upon him, Harvey was at a moment of crisis. The defense of his new thesis of forceful systole had led him into impossibility and threatened the very credibility of the thesis itself. If so much blood passed out of the heart, where did it all go? Why were the arteries not distended with blood? How could the flesh absorb so much nourishment? And where did all this blood come from? Could such a large amount be produced from the ingested food, as Galen said it did? French 1994: 90–91)

As French describes it, Harvey's reasoning is a salient example of incongruity-resolution in that an unexpected observation triggers backward chaining, i.e. reasoning in function of spelling out what made him expect something else than what he observed. However, we would stretch the idea of "forced reinterpretation" too far if we would imply that inferring the contradiction that the veins would be emptied and that the arteries would burst, was purely a matter of backward chaining, for the formulation of the contradiction clearly assumes more than merely Harvey's observation in con-

[15] Related to this is Harvey's alternative explanation for the pulse, but that needn't concern us here.

junction with Galen's account of the production and consumption of blood. One elegant solution to the historiographical problem outlined above would be to suggest that it was precisely Harvey's reconsideration of the function of the venous valves which allowed him to turn the incongruity into a contradiction. After all, on Harvey's account of them, the venous valves only allow for venal blood to flow towards the heart and not also away from it:

> Then suddenly he had the answer. He remembered (he later told Robert Boyle) that the valves in the veins all seemed to point *to* the heart; if they were valves as Harvey understood valves, this meant that the motion of blood in the veins was *to* the heart, *from* the finest ramifications of the veins, where they collected the blood that had been driven to the finest ramifications of the arteries by the forceful systole and pulse. French 1994: 91

In any case, we can quite plausibly apply the explication of incongruity-resolution outlined in Section 2 to this discovery process. What does it mean to say that the observation was unexpected? It simply means that Harvey *could not have predicted* such an amount of blood on the basis of the old theory. Again, just as in the qualification of Suls' two-stage account above, most probably Harvey did not make any precise prediction as to the amount he did expect. In turn, however, the incongruity forced Harvey to reinterpret blood and blood flow, so as to find a way to accommodate his observation, or more precisely: to formulate a theory from which his observation *could have been predicted* (if only he had not been led astray by Galen's long-standing theory).

4. Incongruity-resolution in thought experimentation (2)

According to Ian Hacking (1993), Thomas Kuhn wavered for some time between entitling his seminal (1964) paper "A Function for Thought Experiments" or "The Function for Thought Experiments" . Kuhn eventually opted for the more modest "A". However, as I interpret the paper, one of Kuhn's aims is clearly to advance a specific explication of thought experiments. So he is not merely developing some descriptive account of them, but also making just the kind of prescriptive claims about the proper usage of the term that would legitimate a "The" in the title.

At any rate, the phenomena for which Kuhn reserves the term are initially singled out on the basis of their function. From the outset Kuhn spells out that he will assess critically the prevalent idea that thought experiments in

physics do not produce a new understanding of nature but rather of the scientist's conceptual apparatus. So his paper is, at least initially, about a specific class of thought experiments, namely those that effect *conceptual change*.

In the first two sections, Kuhn seems to agree on most points with that received view. He analyzes both a non-historical and a historical example of the induction of conceptual revision, the former being a laboratory experiment performed by the famous Swiss child psychologist Jean Piaget, the latter one of the many thought experiments Galileo deployed to confront readers with the paradoxes implicit in Aristotle's physical concepts, thereby paving the way for specific conceptual changes that would avoid the signaled difficulties.

In section three, however, Kuhn forcefully criticizes the prevalent view and especially the idea that the old concept, i.e. the concept which the thought experiment forces us to revise, was "self-contradictory" or "confused" all along. I believe it is worth quoting Kuhn's outpouring with respect to the analysis of Galileo's thought experiment in terms of the prevalent view at some length:

> Aristotle, if no experimental physicist, was a brilliant logician. Would he, in a matter so fundamental to his physics, have committed an error so elementary as the one we have attributed to him? Or if he had, would his successors, for almost two millennia, have continued to make the same elementary mistake? Can a logical confusion be all that is involved, and can the function of thought experiments be so trivial as this entire point of view implies? I believe that the answer to all of the questions is no, and that the root of the difficulty is our assumption that, because they rely exclusively upon well-known data, thought experiments can teach nothing about the world. Though the contemporary epistemological vocabulary supplies no truly useful locutions, I want now to argue that from thought experiments most people learn about their concepts and the world together. In learning about the concept of speed Galileo's readers also learn something about how bodies move. What happens to them is very similar to what happens to a man, like Lavoisier, who must assimilate the result of a new unexpected experimental discovery. Kuhn 1964: 320–321

Now, how do these thought experiments work then? At the risk of outrunning Kuhn by using terms he does not employ, I would summarize his account as follows. What a thought experiment actually does is changing "the domain of application" of a concept. Aristotle's concept of motion can be consistently applied to the cases the Aristotelians intended to cover.

Galileo's thought experiment, however, extends that domain of application by considering an unexpected case to which, as it turns out, Aristotle's concept cannot be applied, at least not consistently. So thought experimenters are almost literally caught by surprise: as soon as the thought experimenter accepts that the unexpected case which is put forward is one to which the old concept should be applicable, he or she is not only bound to bump into a difficulty (e.g., an inconsistency), but also to revise the old concept so as to evade the signaled difficulty (e.g., by means of bifurcation).

In the outpouring quoted above, Kuhn protest so vigorously against calling the old concepts which thought experiments force us to revise, "self-contradictory" or "confused", because it is quite natural to embody in our concepts certain expectations about what could happen:

> Ought we demand of our concepts, as we do not and could not of our laws and theories, that they be applicable to any and every situation that might conceivably arise in any possible world? Is it not sufficient to demand of a concept, as we do of a law or theory, that it be unequivocally applicable in every situation which we expect to encounter?
>
> Kuhn 1964: 322

So thought experimentation is yet another example of incongruity-resolution. The old and the revised concept correspond to the two interpretations of the joke boy. The incongruity arises when the novel case is proposed, i.e. a case of which we did not expect that it could happen. At least, that expectation was not build into our concept. Obviously, at this point the thought experimenter has some choice: he or she can decide not to accept the proposed change in the domain of application of the concept. But as soon as the thought experimenter accepts it (and thought experiments are typically designed to make us accept the proposed change in the domain of application), the incongruity forces him or her to reconsider the notion at hand, since, as it turns out, the old concept cannot be consistently applied to the novel case. Moreover, just as in Harvey's case, the backward chaining involved in turning the incongruity into a straightforward inconsistency, seriously facilitates the subsequent forward chaining. As Kuhn expresses it metaphorically:

> Full confusion, however, came only in the thought-experimental situation, and then it came as a prelude to its cure. By transforming felt anomaly to concrete contradiction, the thought experiment informed our subjects what was wrong. That first clear view of the misfit between experience and implicit expectation provided the clues necessary to set the situation right.
>
> Kuhn 1964: 333

However, the application of incongruity-resolution to thought experimentation which I have developed up to now is on the meta-level and not, as in the sections on joke comprehension and scientific discovery, on the object-level. So in order to avoid (further) category-mistakes, I will now discuss an example in which the incongruity arises not by the consideration of an unexpected case as such, but *within* the thought-experimental scenario. Moreover, to drive the point home that Kuhn is really after advancing a specific explication of thought experiments (i.e., not limiting himself to an identifiable subclass) my example will not be a thought experiment that induces conceptual change, but rather one that plays a role in theory choice.[16]

Let me take at once an example which has been analysed *ad nausseam* in the literature on thought experiments. According to Aristotle, heavy bodies fall faster than light ones ($H > L$). Galileo refutes Aristotle's theory by asking us to consider what would happen if a heavy and a light body were strapped together and dropped ($H + L$). On the one hand, the light body will slow down the heavy one, so that the combined system should fall slower than the heavy body alone ($H > H + L$). On the other hand, the weight of the combined system is greater than that of the heavy body alone, so that the combined system should fall faster than the heavy body alone ($H + L > H$). Thus Galileo refutes Aristotle's theory by deriving a contradiction from it. But at the same time, Galileo establishes his own theory, according to which all bodies fall at the same speed regardless of their weight ($H = L = H + L$).

What I have just presented is completely in line with the anachronistic rendering of Galileo's thought experiment in contemporary literature. However, if one actually looks at Galileo's text, a different picture emerges. For sure, the upshot is the same, but we do not get there so easily, i.e. by such a straightforward *reductio ad absurdum.*

To evaluate Aristotle's theory that heavy bodies fall faster, Galileo has us considering what would happen if a heavy and a light body were strapped together.[17] Taken together, the targeted theory, as Galileo describes it, and the novel case he puts to the fore, constitute, so to speak, the set-up of the thought experiment. Subsequently, Galileo has us admitting that the falling

[16]The distinction is drawn carefully by Gendler (2000: 25): whereas "factive thought experiments" ask us what would happen given some imaginary scenario, "conceptual thought experiments" ask us how we should describe what would happen.

[17]On the meta-level, this could already be described as an incongruity, for Aristotle's theory was only meant to account for "natural occurrences". So the case which this thought experiment brings to the fore is "unexpected" as such and, as soon as it is accepted, it changes the original domain of application, not of the concept, but rather of the theory.

speed of the light body will slow down that of the heavy one. This is not what we *expect*, but Galileo brings us to see that we should *accept* it. This proposition acts, so to speak, as the punch line of the thought experiment.

Fortunately, however, Galileo is helpful enough to tell us how the incongruity should be resolved. At the very least, he explicitly helps us in *backtracking*: why did the acknowledgment that the light body would slow down the heavy one *defy our expectations*, because it is at pars with the "stipulation" that heavy bodies fall more quickly than light ones.

So what I have been suggesting here is that incongruity-resolution is the problem-solving process involved in comprehending a thought-experimental *reductio ad absurdum* like Galileo's. Note that the fact that Galileo could develop the same case the other way around, i.e. by first having us *accepting* that the combined system would fall quicker, is not a fatal problem for my analysis. For the thought experiment would only be rhetorically effective to the extent that accepting that proposition would *defy expectations* induced before, i.e. in what one would call the "set-up" of the thought experiment.

5. Conclusion and further research

Taking as my starting point a tentative and partial explication of "incongruity-resolution" as it occurs in humor comprehension, I have argued in this paper that similar reasoning processes are involved in discovery processes which are triggered by unexpected observations, such as Harvey's discovery of blood circulation and Lavoisier's discovery of oxygen, and in thinking through the kind of thought experiments Thomas Kuhn (1964) has tried to single out in his paper "A Function for Thought Experiments". The theoretical benefits of identifying a single problem-solving process occurring in contexts as diverse as humor comprehension, scientific discovery and thought experimentation are obvious, for as soon as one has developed an *explication* of the reasoning processes involved in one of these domains, only minor modifications will be required for it to be successfully applied to the other two domains as well. However, in none of the domains mentioned, a satisfactory explication has been advanced yet. Therefore I will briefly speculate in this concluding section about the lines along which such an explication of incongruity-resolution might be developed.

Most scholars will agree that the reasoning processes involved in humor comprehension, scientific discovery and thought experimentation exhibit an external dynamics, i.e. quite often, a conclusion is withdrawn in view of new information. So a first constraint on the envisaged explication of incongruity-

resolution is that it should do justice to the fact that the derivability relation is typically *non-monotonic*. Now there are some logics on the market which allow for that. There is but one kind of logics, however, which can also do justice to the *internal dynamics* involved in incongruity-resolution, i.e. when a conclusion is withdrawn in view of the better understanding of the premises provided by a continuation of the reasoning. That kind of logics is called adaptive logics and they have been developed over the last two or three decades to explicate, by means of dynamic proof theories, many interesting dynamic consequence relations, like the reasoning processes involved in induction, abduction, metaphors and scientific discovery.[18]

What an adaptive logic does is supposing that all formulas behave normally *unless and until* an abnormality arises. What an abnormality is, depends on the logic at hand: many of them handle inconsistencies, but there are also adaptive logics which deal with, e.g., ambiguities. Importantly, if such an abnormality occurs, it is considered to be *local*, i.e. the problem only affects the formula behaving abnormally and its consequences, but not at all the formulas behaving normally and their consequences. Now it is clear that the explication of incongruity-resolution could deploy the tools developed within this adaptive logic-program. For one thing, incongruities are patently *local* problems or "abnormalities". For another thing, the backtracking postulated by the forced reinterpretation-account of incongruity-resolution, clearly exhibits internal dynamics.

This being said, one can also anticipate certain problems which will arise when one tries to develop, say, incongruity-adaptive logics. E.g., on the present analysis, incongruity-recognition, i.e. the ability to "see" that the punchline and the initial interpretation are incongruous, without having derived, e.g., an inconsistency, is something extra-logical. Although the outlined explication of the reasoning process can, at least in principle, cover the inferences made during backward chaining, it doesn't sufficiently pinpoint what triggers that backward chaining in the first place. Invoking extra-logical elements is of course somehow disappointing. But it might turn out to be irremediable. At face value, at least, explications that do not invoke backward chaining at all, would also depend on extra-logical elements, for they need to invoke the ability to "see" the opportunity to take something humorously. One might even say that on an analysis solely in terms of for-

[18] For more on the adaptive logic-program as it is developed in the Centre for Logic and Philosophy of Science (Ghent University), see its homepage at http://logica.ugent.be/adlog/al.html . For adaptive logics and discovery processes, see especially Meheus (1993; 1999) and Meheus & Batens (1996).

ward chaining, there could not be any difference in principle between "seeing the opportunity to take something humorously" and "grasping something as humorous", so that, at the end of the day, "humor comprehension" would be degraded from a genuine problem-solving process to a mere problem-defining activity.

In any case, because it doesn't restrict itself to capturing the external dynamics of reasoning processes, the adaptive logic-program promises to supplement the forced reinterpretation-account of incongruity-resolution with the tools necessary to construct dynamic proofs, which do not mimic, but rather *explicate* how "new ideas" are shaped during backtracking. Such an explication would not only allow us to show how backward chaining can facilitate subsequent forward chaining dramatically, it would also provide us with the means to express, in a more formal way, the resemblances between humor comprehension, scientific discovery and thought experimentation to which both Arthur Koestler and Thomas Kuhn already alluded in 1964.

References

Attardo, S. (1994), *Linguistic Theories of Humor*. Berlin, New York: Mouton de Gruyter.

Attardo, S., and V. Raskin (1991), "Script theory revis(it)ed: joke similarity and joke representation model", *Humor: International Journal of Humor Research* 4, 293–347.

Brône, G., and K. Feyaerts (2003), "The cognitive linguistics of incongruity resolution: marked reference-point structures in humor". Preprint 2005, Department of Linguistics, KULeuven.

Bruner, J., and L. Postman (1949), "On the perception of incongruity: a paradigm", *Journal of Personality* 18, 206–223.

Buckley, F. H. (2003), *The Morality of Laughter*. Ann Arbor: The University of Michigan Press.

Bylebyl, J. J. (1982), "Boyle and Harvey on the valves in the veins", *Bulletin of the History of Medicine* 65, 351–367.

De Mey, T. (2003), *Thinking Through Thought Experiments*. Unpublished dissertation, UGent.

French, R. (1994), *William Harvey's natural philosophy*. Cambridge, New York, Melbourne: Cambridge University Press.

Gendler, T. (2000), *Thought Experiment: On the Powers and Limits of Imaginary Cases*. New York, London: Garland Publishing.

Hacking, I. (1993), "Do thought experiments have a life of their own?". In: *Proceedings of the Philosophy of Science Association*, vol. 2, pp. 302–308.

Harvey, W. (1578), *The Movement of the Heart and Blood*. Translated with introduction and notes by G. Whitteridge (1976). Oxford, London, Edinburgh, Melbourne: Blackwell Scientific Publications.

Kuhn, T. (1962/1970), *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Kuhn, T. (1962), "The historical structure of scientific discovery", *Science* 136, 760–764.

Kuhn, T. (1964), "A function for thought experiments". In: *Melanges Alexandre Koyre*. Paris: Hermann, pp. 307–334.

Langeley, P., H. A. Simon, , G. L. Bradshaw, and J. M. Zytkow (1987), *Scientific Discovery: Computational Exploration of the Creative Process*. Cambridge: MIT Press.

Morreall, J. (1987), *The Philosophy of Laughter and Humor*. Albany, NY: State University of New York.

Meheus, J. (1993), "Adaptive logic in scientific discovery: The case of Clausius", *Logique et Analyse* 143–144, 359–391.

Meheus, J. (1999), "Clausius' discovery of the first two laws of thermodynamics: a paradigm of reasoning from inconsistencies", *Philosophica* 63, 89–117.

Meheus, J., and D. Batens (1996), "Steering problem solving between cliff incoherence and cliff solitude", *Philosophica* 58, 153–187.

Nickles. T. (1980), *Scientific Discovery, Logic, and Rationality*. Dordrecht: Reidel.

Pagel, W. (1976), *New Light on William Harvey*. Basel, New York: S. Karger.

Raskin, Victor (1985), *Semantic Mechanisms of Humor*. Dordrecht, Boston, Lancaster: D. Reidel.

Ritchie, G. (1999), "Developing the incongruity-resolution theory". In: *Proceedings of the AISB Symposium on Creative Language*. Edinburgh, Scotland, pp. 78–85.

Ritchie, G. (2004) *The Linguistic Analysis of Jokes*. London, New York: Routledge.

Simon, H. A. (1977), *Models of Discovery*. Dordrecht: Reidel.

Suls, J. (1972), "A two-stage model for the appreciation of jokes and cartoons". In: J. H. Goldstein and P. E. McGhee (eds.), *The Psychology of Humor*. London, New York: Academic Press, pp. 81–100.

Suls, J. (1983), "Cognitive processes in humor appreciation". In: P. E. McGhee and J. H. Goldstein (eds.), *Handbook of Humor*, Research. New York, Berlin, Heidelberg, Tokyo: Springer, pp. 39–57.

Veale, T. (2004), "Incongruity in humor: Root cause or epiphenomenon?", *Humor* 17, 419–428.

Whitteridge, G. (1971) *William Harvey and the Circulation of the Blood*. London, New York: Macdonald, American Elsevier.

TIM DE MEY
Centre for Logic and Philosophy of Science
Ghent University
Blandijnberg 2
B-9000 Gent, Belgium