Creativity, Humour, and Cognition

Mario Gensollen
UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES
mgenso@correo.uaa.mx

Marc Jiménez-Rolland
UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES
marcji2121@yahoo.com

ABSTRACT
This paper explores some aspects of the scientific study of creativity by focusing on intentional attempts to create instances of linguistic humour. We argue that this sort of creativity can be accounted for within an influential cognitive approach but that said framework is not a recipe for producing novel instances of humour and may even preclude them. We start by identifying three great puzzles that arise when trying to pin down the core traits of creativity, and some of the ways taken by Cognitive Studies in this quest. We then consider what we call 'creative humour', which exhibits the core features of the aforesaid creativity. We then explore how a key cognitive approach to human communication can account for creative humour. We end by drawing lessons and highlighting limitations to cognitive approaches to creativity.

Keywords: linguistic humour, creative humour, incongruity-resolution, relevance theory   .

INTRODUCTION

In 'The Hesitation Ramification', an episode of the popular sitcom *The Big Bang Theory,* Dr. Sheldon Cooper is searching for a Unified Theory of Comedy. He deems this pursuit a mere intellectual curiosity, since he describes himself as 'hysterical'. Envisaging such a theory, one can imagine it would not only provide a deeper understanding of human humour,

but it would yield the advantage (especially important for someone like Sheldon) of allowing one "…to elicit laughter from anyone at any time (unless they're German, 'cause that's a tough crowd)". As it is often the case in the show, one is left wondering exactly how Dr. Cooper's endeavours are doomed to fail, whether due to a lack of ability, a flaw in execution or even whether the goal itself is a wild-goose chase.

One can readily accept that a unified theory is not a precondition for creating humour. The main question we address in this paper is whether developing such a theory would help in creating humour.

In broad terms, this paper explores a paradox concerning the scientific study of creativity. The result is an outline of the 'paradox for a theory of creativity'. A scientific theory should convey information about the phenomena it covers. If there is a theory, it follows that the phenomena are not novel. Furthermore, it is widely believed that an explanatory theory should render its outcomes predictable. Put another way, its outcomes should be unsurprising in the light of the scientific explanations given. Yet both novelty and unpredictability seem to be the hallmarks of creativity. It would thus seem that creativity cannot possibly be explained by a scientific theory. Although we deem some aspects of the paradox to be grounded on misconceptions about scientific theories, here we stress those linked to creativity. To tackle these issues, we focus on a narrower target by addressing attempts to intentionally produce creative instances of linguistic humour. If one can find a theory explaining this kind of creativity, there may be hope for a full-blooded cognitive explanation of overall creativity. We argue that much of what is interesting about humorous creativity can be accounted for within the framework of an influential cognitive approach and by producing algorithmic inference patterns (whether conscious or not). This, however, does not mean that such an approach will "elicit laughter from anyone at any time." It is highly unlikely that anyone will hit upon "a fool-proof recipe for generating humor stimuli of all varieties" (Hurley *et al*., 2011, p. x). Indeed, a theory on humorous effects could even preclude them because everyone knows a common way to kill a joke is to explain it.[1] We therefore try to show that the 'paradox for a theory of creativity' is based on an illusion that — like The Cheshire Cat — ends up vanishing.

To support our thesis, we will first identify three great puzzles that arise when trying to naturalise creativity. After explaining each of these puzzles, we outline ways in which cognitive studies have tried to solve them. In a second section, we focus on the links between creativity and humour. Although there is much of empirical interest on this score, we shall confine discussion to what we call 'creative humour'. After a ready-made characterisation, we will show how creative humour displays the core features of creativity identified earlier. The third section shows how a major cognitive approach to human communication can account for creative humour. Although cognitive studies of creativity have much to offer, we end by drawing lessons and highlighting some limitations for a robust naturalisation of creativity as a phenomenon for enquiry by the discipline.

NATURALISING CREATIVITY
Creativity is one of Man's hallmarks as a species. We engage in creative tasks in many key areas such as Science, Engineering, Art, Design, Gastronomy, Entertainment, Sports, and a host of other practices and activities, in which the term 'creative' is applied to agents, processes, and products. Moreover, creativity appears to be a solely human trait; when it is extrapolated to inert physical processes, there seems to be an assumption that rational agency is somehow involved. Although acknowledging its special features, a successful naturalisation of creativity would yield an understanding akin to that found in scientific scrutiny in other fields. To gain such understanding, we must be able to show that creativity involves physical entities and mechanisms that can be embodied within a systematic theory. However, there seem to be special hurdles to recognising creativity as a phenomenon within the natural order. We go on to identify three such hurdles to naturalizing creativity: (1) *semantic diversity; (2) normativity*; (3) *unpredictability*. After discussing them, we outline some ways in which these hurdles might be overcome.

---

1   Thus, one could agree with the *New Yorker*'s cartoonist Robert Mankoff, who — retrieving E. B. White quip — points out that "analyzing humor is like dissecting a frog. Few people are interested, and the frog dies of it" (2009, p. ix).

The first challenge for coming up with a scientific theory of creativity is to pin down what the subject matter of such a theory should be. The term 'creativity' is associated with a vast array of meanings. Being creative seems to involve fluency, flexibility, divergent thinking, innovation, discovery, originality, spontaneity, genius, wit, ingenuity, shrewdness, imagination, insight, acumen, and so forth. It can be argued that 'creativity' belongs to a large semantic family with countless relatives. Thus, the term 'creativity' seems to exhibit a semantic diversity that makes it ill-suited for scientific theorising. As Stokes and Paul pointed out:

> We find creativity not only in art but also in science, theorizing of any sort, engineering, business, medicine, sport, gaming, and so on. At least two worries may be developed accordingly. First, given the complexity of any one of these individual domains, we might worry that there are simply too many variables to allow for a clear explanation (…). The second worry concerns generalizability. Even supposing that we could explain the creative achievement of some artistic master, any such explanation will have to be so specific that it will fail to generalize to artistic creativity, or creativity in other domains like science or gaming or whatever. In short, given the variety of creativity (and the complexity of the varied creative achievements), identifying a general explanation, in the form of a set of cognitive and behavioral features, may seem entirely improbable (2016, p. 320).

Although this might seem a case of rampant ambiguity, there is at least one constant element in most conceptions of creativity, namely, they all appear to involve some kind of *novelty* (Cropley, 2011, p. 511). However, being novel is a necessary but not a sufficient condition for being creative.

Normativity constitutes the second challenge to drawing up a scientific theory of creativity. Let us consider the case of novelty without creativity. This can be illustrated by considering drug patents. In some countries, drug prices can only legally be raised if pharmaceutical companies can prove that the drug has been modified, thus justifying the need to recoup the cost for the development of new drugs. Thus, drug patents can be extended so long as companies make changes to the product. Yet many observers consider companies abuse the system, making trifling changes to existing drugs just so they can re-file the patents — a practice known in the industry as 'evergreening' and 'product hoping' (Ward, Hickey and Richards, 2020, pp. 1-2, 19-24). Legally, the companies do not have to prove that the new drug is any better than the old formulation, merely that it is not worse. Although such cosmetic modification is legal, it is clearly not creative. This is why some argue that companies should not be allowed to extend their patent rights by using such a ruse given that novelty without value is hardly creative. This raises the question as to why we value *some* novel ideas but not other. What is valuable may vary across the spectrum of human interests. Naturalising the value of creativity involves at least three steps. The first step is to identify our goals. The second is to show how those goals are useful. The third is to explain how 'creative' achievements meet those goals. On these lines, full naturalisation of normativity would require the properties expressed by evaluative or normative terms (in this case, the term 'creative' to be natural or to depend upon natural properties).

The third hurdle is that creativity does not seem to fit into the natural order of things. For something to be creative, it is usually suggested that it should not be caused or conditioned. A common way to spell out this apparent feature of creativity is to invoke the quasi-esoteric acceptance of an inexplicable source of inspiration (we can trace the lineage of this idea back to Plato). However, we think that this way of understanding the nature of the challenge can lead to confusion. That is because it relies on the assumption that a naturalistic theory should always have predictive power. If it fails to provide accurate predictions, non-natural explanations may be at fault. Another way to state the problem recognises that the phenomena dealt with by a theory of creativity should somehow be *unpredictable*. This challenge

acknowledges that, while it is quite possible that a theory lacks predictive ability, it may still provide *bona fide* naturalistic explanations (as, for example, does Evolutionary Biology). The challenge lies in demonstrating such theories' robust explanatory power while saying why they do not serve for making predictions.

There are some promising ways in which these challenges have been addressed by cognitive studies of creativity. Regarding the first one, it has been argued that the concept of creativity is an *integrative* or *inclusive* one. The complexity and polysemy of the concept imply that when we address different perspectives, we find that they overlap and interact with one another. If the concept of *creativity* were not an inclusive one, we could not refer to it as a conceptual category (Estany and Herrera, 2016, p. 96).

With regard to the normativity hurdle, several attempts at naturalisation have been made. Nevertheless, creative instances could only be assessed in relation to a vast array of human interests. Why is a creative instance valuable? Why do we seem to value creativity in general? This second issue may deepen our knowledge on the first one. Given the complexity of individual domains, creative instances may lead us to conclude that there are simply too many variables to allow us to come up with a clear explanation. Nonetheless, we can find a promising path from novelty to normativity by drawing a distinction between someone who is psychologically creative (P-creative) and someone who is historically creative (H-creative): "…some people repeatedly produce ideas highly regarded as valuable — and which, so far as is known, no one else has ever had before (…) Most people, by contrast, produce only moderately interesting ideas, many of which are already known by other people" (Boden, 2009, p. 237). Although, what we regard as interesting and valuable varies, there are many different ways for something to be praised as valuable, and historical creativity is very hard to gauge. One way of unravelling this problem is to first understand psychological creativity (for this is at least necessary for historical creativity)

and then understand how someone comes up with a wholly new idea.[2]

Regarding the third hurdle, it has been argued that creativity is compatible with determinism and therefore with naturalistic explanation (Kronfeldner, 2009). Psychological creativity seems to involve some kind of originality and spontaneity and thus seems to be independent from social learning, experience and prior knowledge. Nonetheless, as Kronfeldner shows, this independence is compatible with determinism. While creativity appears to be opposed to specific causal factors, it does not exclude causal determination. Thus, as we pointed out earlier, the third hurdle is not an insurmountable one. Additionally, Boden (2009) has pointed out that different kinds of creativity operate within the framework of a shared conceptual base or within conceptual spaces and thus creativity is not at odds with restrictions. Moreover, as Boden suggests "…exploratory creativity, the existing stylistic rules or conventions are used to generate novel structures (ideas), whose possibility may or may not have been realised before the exploration took place" (2009, p. 241).

As we saw above, although novelty is a recurring element in the various definitions of creativity, seizing on a given example does not prove the link. Fortunately, Boden provides a useful definition that allows us to capture the main features mentioned so far, without raising further issues on the aforementioned three hurdles. She defines creativity as "…the ability to come up with ideas that are *new*, *surprising*, and *valuable*" (Boden, 2004, p. 1). The label 'ideas' is meant to be a catch-all term covering a vast array of feats, including poetic images, scientific theories, works of art, culinary dishes, design solutions, and

---

2 As one of the reviewers for this paper pointed out to us, coming to terms with this aspect of the puzzle of normativity requires taking into account several key components of the social dimension. While we do not underestimate the significance of these steps towards an overall scientific understanding of creativity, further remarks on this inquiry into the depths of historical creativity are beyond the scope of this paper.

winning strategies. Explaining creativity under this definition would involve surmounting the hurdles to solve what we term 'the paradox for a theory of creativity'. This definition could also be useful for assessing apparent instances of creativity and — as we will see below — will let us draw distinctions between humour in general and creative humour in particular.

CREATIVITY AND HUMOUR

Research on the links between humour and creativity has been undertaken from neurological, psychological, cognitive and philosophical standpoints. Some recent studies have explored the neural correlates of creativity and linked them to humour. The creative tasks investigated commonly range from narrative generation (Howard-Jones *et al.*, 2015) to jazz improvisation (Limb and Braun, 2008). However, the researchers unfortunately discovered that the cortical regions associated with creativity were not linked and varied depending on the activity involved. Yet Dietrich and Kanso (2010) observed the common involvement of the prefrontal cortex. Meanwhile, Amir and Bierderman (2016) argued that a one-dimensional comparison between creative and non-creative control conditions might be ill-suited for showing the brain functions involved in creative ventures. Furthermore, for them the neural correlates of real-time humour creation had been too little explored. They showed through neuro-imaging that greater comedic experience is associated with less activation in the stratum and medial pre-frontal cortex but with greater activation in the temporal association regions.

Psychologists and cognitive scientists had found several links and some correlations between humour and creativity since the 1960s. Getzels and Jackson (1962) studied how highly creative groups made more use and valued humour more than non-creative ones. Gordon (1962) showed that people involved in developing creative problem-solving systems reported that sense of humour is a trait consistently present

in trainees who are comfortable in dealing with analogies and associative forms of thought. Treadwell measured the ability to create humour and related that ability to other measures of creativity, and concluded that the "…study of humor appears likely to be a useful approach in the study of creativity" (1970, p. 57). For a variety of researchers, humour can be considered an aspect of creativity (Amabile, 1987; Arieti, 1976). In fact, the way researchers measure creative skills usually includes valuations of humour (Davis and Subkoviak, 1975; Torrance, 1966). Others considered humour as creative expression (Koestler, 1964; Maslow, 1971). For Murdoch and Ganim, "… humor seemed to be sufficiently integrated to be considered a subset of creativity", and that is why "…the two could be productively studied within similar conceptual frameworks" (1993, p. 66). Some researchers tend to treat sense of humour as a positive trait (Beermann and Ruch, 2009; Hong, 2010). For Ziv (1976) a humour-filled atmosphere promotes creative performance and people who were required to apply their sense of humour commonly use non-traditional thinking, which enhances their creativity. The Hungarian-British author and journalist Arthur Koestler (1964) posited a strong correlation between humour and creativity: for him humour, scientific discovery, and artistic creation are forms of creativity that all involve the same cognitive process, which Koestler called 'bisociation'. He coined this term "… in order to make a distinction between the routine skills of thinking on a single 'plane', as it were, and the creative act, which (…) always operates on more than one plane" (Koestler, 164, pp. 35-36).

There is another link between humour and creativity that has also prompted research interest, namely that humour can be an instance of creativity. Thus Chan, Chen and Lavallee have pointed out that "… humor not only facilitates creativity but may also be a display of creativity in and of itself" (2013, p. 610). Not all instances of humour are creative but some are. As we saw in the previous section, creativity seems to involve novelty, surprise and value. So, if some instances of humour involve creativity, they would need to be at least novel, surprising and valuable.

Theoretical approaches to humour — whether creative or not — have a long history. Philosophers took an interest in humour from the outset. Several philosophers (such as Plato, Aristotle and Hobbes) gave us insights on the *laughing-at* phenomenon. Superiority Theories (STs) of humour were mostly looking for the psychological causes of laughter and amusement: advocates of STs "…said that when something evokes laughter, it is by revealing someone's inferiority to the person laughing" (Morreall, 2009, p. 7). A vast quantity of instances of humour fit well under STs: "We often laugh *at people*. The implied superiority is what explains the well-worn excuse: I'm not laughing *at* you; I'm laughing *with* you" (Hurley *et. al.*, 2011, p. 41). Here is a cruel joke about lawyers that exemplifies STs:

> Four surgeons were taking a coffee break and were discussing their work. The first said, "I think accountants are the easiest to operate on. You open them up and everything inside is numbered".
>
> The second said, "I think librarians are the easiest to operate on. You open them up and everything inside is in alphabetical order".
>
> The third said, "I like to operate on electricians. You open them up and everything inside is color-coded".
>
> The fourth one said, "I like to operate on lawyers. They're heartless, spineless, gutless, and their heads and their asses are interchangeable" (Hurley et. al., 2011, p. 41).

STs are searching for a causal explanation, not for a teleological explanation, a cognitive account, or a conceptual analysis (although some specific STs may cover some of these aspects). For example, Plato thought that laughter elicited by humour targets a given vice: unawareness. We laugh at people who do not know themselves and think that they are better than they really are. It is in this sense that laughter can be understood as a form of *abuse*. Similarly, Roger Scruton (1982) sees the laughing-at phenomenon as a device for the devaluation of the object of laughter in the subject's eyes. As far as Plato was concerned,

laughter had no place in a well-ordered society because it undermined co-operation and tolerance. He also thought that laughter overrode rational self-control. His pupil Aristotle partially shared this perspective: he defined humour as a form of abuse and conjectured that comedy began as invective (Carroll, 2014, p. 6). In brief, for Aristotle "…humor is the recognition of a failing or a piece of ugliness, resulting from an implied comparison between a noble state of a person or thing and an ignoble state" (Hurley *et. al.*, 2011, p. 41). STs can also explain cases in which one laughs at oneself. Thomas Hobbes, the paradigmatic defender of STs, famously remarked: "*Sudden Glory*, is the passion which maketh those *Grimaces* called Laughter; and is caused either by some sudden act of their own (…); or by the apprehension of some deformed thing in another." (1651, p. 43). Therefore, when we laugh at ourselves, "…we do so putatively from a present perspective of superior insight that sees and savours the ridiculous absentmindedness of the person we once were" (Carroll, 2014, p. 9). Finally, Henri Bergson, drawing upon STs, saw laughter as a social corrective.

STs face major challenges and limitations, namely: (a) feelings of superiority are not a necessary condition for laughter; (b) the recognition of our superiority to others does appear be a sufficient condition for laughter (as Francis Hutcheson memorably pointed out, we realise that we are superior to oysters but we do not laugh at them); (c) we can laugh at comic characters superior to us; (d) it is hard to explain in terms of feelings of superiority why we laugh when we are teased in a friendly fashion; and (e) often the source of laughter has nothing to do with issues of superiority and inferiority (Carroll, 2014, pp. 8-16). Consider the following joke: "Theater sign typo: Ushers will eat latecomers" (Hurley *et. al.*, 2011, p. 41). Jokes like this are very hard to explain under STs. Additionally — as Hurley, Dennett and Adams suggested — this account of humour faces a core weakness: "although it provides a generic reason underlying much (if not all) humor, it does not provide a *mechanism* of humor, and thus also doesn't provide a reason for the reason!" (2011, p.

42). Nonetheless, STs have the virtue of covering many humoristic instances: for example those that make fun of foolishness. STs can also cover the value of humour pointing out that laughter is pleasant, and the pleasure we feel is elicited by the recognition of our actual superiority over the object of laughter.

Release Theories (RTs) of humour mostly looked at the value of laughter and comic amusement. Why do we spend so much time and money consuming humour-based products? Noting the purpose of humour, RTs claim that what fulfils this purpose is what we call humorous. Prompted by the Earl of Shaftesbury, Freud and Spencer (and maybe Aristotle in the lost second book of his *Poetics*), what RTs stress is "…that tension from thought can build up, and when this tension is released by a positive emotion that results from further thought, the energy is transformed into (or spent by) laughing" (Hurley *et. al*., 2011, p. 44). For Shaftesbury, the natural free spirits of ingenious men will find out other ways to slip from their constraints and revenge themselves on those who constrain them (Morreall, 2009, p. 16). In Freud's version, "…certain events create repressed sexual and/or aggressive energy, and when that tension is undone in a dramatic way (suddenly or surprisingly), rather than gradually, the nervous energy is released, and relief ensues in the form of humor" (Hurley *et. al*., 2011, p. 44). Despite the appeal of RTs, which explain the prominence of sexual and aggressive content in humour, RTs cannot explain logical humour. Simple puns and grammatical traps do not necessarily include aggressive or sexual tension.

Incongruity-Resolution Theories (I-RTs) of humour mostly looked for the mechanism(s) eliciting laughter and comic amusement. Strongly championed by psychologists, philosophers and cognitive scientists, I-RTs tell us that "…humor happens whenever an incongruity occurs that is subsequently resolved" (Hurley *et. al*., 2011, p. 45). Also, incongruity is a relational notion: "It presupposes that something is discordant with something else. When it comes to comic amusement, that *something* else is how the world is or should be" (Carroll, 2014, p. 18). A classic example would be this one:

> 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?" (Hurley *et. al*., 2011, p. 46).

I-RTs explains comic amusement elicited by this joke pointing out that O'Riley response is incongruous with being found not guilty.

Let see another one:

> A somewhat heavy man goes into a pizza parlour and orders a pie. The man behind the counter says, "Do you want it cut into eight slices or four?" He thinks for a second and says, "Well, four. I'm on a diet".

I-RTs explain the comic amusement elicited by this joke pointing out that our common heuristics can go wrong in certain circumstances. Some I-RTs could add that the value of this humorous instance lies in highlighting a glitch in the way we think — something that helps our cognitive wellbeing.

Championed by Hutcheson, Kant and Kierkegaard, I-RTs have many advantages. They have practical use value: they provide us "…with a useful heuristic for the future comic research by guiding us toward the kinds of variables we should attend to when investigating specimens of invented humour such as comic narratives", and "…with an eminently serviceable method for discovering the secret to the humor one encounters daily in the form of jokes, comic asides, cartoons, sitcoms, and so on" (Carroll, 2014, p. 2).

However, it is unlikely that STs, RTs and I-RTs can fully capture the nature of humour. Nonetheless, we agree with Carroll in that "…using the incongruity theory as a heuristic may pave the way for superior successor theories" (Carroll, 2014, p. 2).

For our current purposes, we can work from a characterisation like this: comic amusement is an emotion that is aimed at particular objects, such as linguistic and intentional jokes, which meet the

criteria posed by I-RTs, where such appraisals then lead to enjoyment and a sense of levity which itself correlates with increased activation of the reward network in the brain's limbic system. The general name for all those objects that give rise to comic amusement is humour.[3]

We shall focus on linguistic creative humour. Specifically, we address *linguistic* creative humour as the production of verbal *stimuli* that set out to cause comic amusement. In creative humour, novelty, surprise, and even value, could be related to cognitive aspects of the producers and consumers of humour. A humorous instance of creativity seems to require the search for solutions linking disparities in an original way that sparks surprise (Kellner and Benedek, 2016; O'Quin and Derks, 2011; Rouff, 1975). Therefore, the humorous novelty could be understood as a type of *combinational creativity* where familiar ideas are combined in unknown ways (Boden 2004, p. 3; 2009, p. 240; 2016, p. 68). The new humorous combination elicits, as Boden points out, a statistical surprise whose object is what was previously considered unlikely. Yet even this improbability is intelligible, therefore valuable. As Boden concludes, the value of a creative instance depends on judgements of relevance (2016, p. 68). It is important to take into account that, in several cases, humorous effects ('finding something funny') involve some element of surprise: incongruity resolution; hence, creative humour, insofar as it requires bringing about incongruence, involves 'crafting surprise'. It follows that the paradox of a theory of creativity would disappear if the creative humour is not merely based on repetition of a social model or habit). This is so because such creative humour would clearly involve the elements ofnovelty, surprise and value.

COGNITIVE MODELS OF CREATIVE HUMOUR
We are now in a position to frame our initial question concerning the prospects of naturalising creativity in a more bounded way. From now on, our concerns will be confined to the main manifestation of creativity that we labelled 'creative humour' in the foregoing section. We will explore to what extent this kind of creativity has been brought within the scope of a systematic understanding, involving physical entities that exhibit it through recognisably natural mechanisms. Several (perhaps complementary) approaches from psychology have tackled various aspects of humour production (relating to personality, social interactions, developmental stages, abnormal behaviour, and so forth)[4]. Here though, we shall focus on theories of the cognitive aspects of creative humour. Furthermore, the range of humour output we consider is confined to that conveyed by language. To assess and support our main contention (namely that that linguistic creative humour can be accounted for within a cognitive framework) we shall first present the main general tenets of Relevance Theory [RT] (Sperber and Wilson, 1987; 1995; Wilson and Sperber, 2004). We then go on to outline how the theory has been applied to research humour in general and the intentional production of linguistic humour in particular. We note some of RT's accomplishments, promising avenues, and limitations.

Relevance Theory as a cognitive account of human communication
Although it is not the only cognitive approach currently on the market, RT has much to offer as a theoretical framework. According to Sperber and Wilson, RT's "aim is to identify underlying mechanisms, rooted in human psychology, which explain how humans communicate with one another" (1995, p. 32). RT seeks to provide an empirical psychological theory that can account for human communication and cognition, making use of some assumptions from (and searching to achieve integration with) Evolutionary Biology. It presupposes that "human cognitive abilities are a part of nature; (…) adapted as a result of natural evolution" (Sperber and

---

3   We take the general lines of this characterisation from Carroll (2014, p. 5).

4   For a contemporary overview, see Gibson (2019).

Wilson, 1995, pp. 116-117). As part of our evolutionary endowment, RT posits the existence of a cognitive mechanism (which might not be a solely human one) that focuses our attention on what might be relevant but that makes us ignore stimuli which are not.

It is worth noting that, by claiming that human "… cognition is a biological function" which "tends to be geared to the maximisation of relevance", RT does not seek to provide a full description of its "process of Darwinian natural selection (or other evolutionary forces that may have helped to shape it)" (Sperber and Wilson, 1995, p. 261). Instead, it is mainly deployed as a theory at the functional or computational level of explanation (Sprevak, 2016, §4; Bermúdez, 2006; 2014). Put another way, it purports to describe the tasks performed by an organism and their ecological purpose. It may also shed light on the algorithmic descriptions of the cognitive processes at work.[5]

The main mechanism posited by RT is called "cognitive principle of relevance". It depicts the human cognitive system as wired to look for relevance and dismiss irrelevance, being able to (unconsciously) rank different processing outputs (e.g., interpretations) produced by the same stimuli. The notion of 'relevance' here is a technical one. It can be applied both to: (1) external stimuli and; (2) internal representations arising from an individual's cognitive processing of an input at a given time, and in a context that yields background information that is expressed in the form of his assumptions. Thus "it makes no sense to talk about the relevance of an ostensive stimulus on its own", instead "relevance is a notion relative to an individual, in a particular context, at a particular time" (Curcó, 1997, p. 169). In addition, there are degrees of relevance and it all boils down to maximising benefits (positive

cognitive effects) while minimising costs (mental processing load).[6]

Cognitive relevance presumably plays a crucial role in achieving communication. This does not require a previously shared code, though it may use one. Rather, it can be seen as an attempt to convey someone's informative aims followed by successful recognition of the message by the target audience (i.e. the communicative goal is achieved).[7] RT holds that ostensive communication requires that the communicator provide an ostensive *stimulus* (i.e., direct evidence of his intention to provide information). This might "create precise and predictable expectations of relevance not raised by other inputs" (Wilson and Sperber, 2004, p. 611), triggered by the so-called 'communicative principle of relevance', which states that "every ostensive stimulus conveys a presumption of its own optimal relevance" (Wilson and Sperber, 2004, p. 612). The communicator has the further task of making "…correct assumptions about the codes and contextual information that the audience will have access to and be likely to use in the comprehension process (…) so that all the hearer has to do is go ahead and use whatever code and contextual information come most easily to hand" (Sperber and Wilson, 1995, p. 43).

Under this view, even if communication involves an elaborated, conventionally shared code such as a natural language, there are significant gaps between the communicator's overt ostensive stimulus (the speaker's

---

5 Thus, it could be argued that RT is also (aiming for) an explanation at the algorithmic level, insofar as there is also an attempt to codify, in a finite number of steps, how the organism can perform the task. However, thus far at least, an exploration of the physical changes in the organism corresponding to those algorithmic steps —what would be the implementation level of explanation— has not yet been pursued or even envisaged.

6 Although our presentation of the cognitive principle of relevance is spelled out in terms of costs and benefits, assumptions and hypotheses, the main tenets of RT do not require this framing in terms of rational choice theory [RCT]. Interestingly, the relevance-theoretic framework can manage without conceiving inference from such an individualistic approach. We thank one of the anonymous referees for pointing this out. However, for expository purposes, in what follows we use the RCT lingo for an expedient exposition.

7 As many naturalistic theories of linguistic meaning that assume that intentionality of thought is explanatorily prior to that of language, insofar as they "take mental representation to be basic and linguistic representation to be derivative", RT assumes that representation in public language is (partially) explained by "the representational powers of mental states" (Papineau, 2006, p. 175).

utterance) and what he intends to communicate ('what he means'). "These gaps are filled by inference" (Yus, 2016, p. xvi). Although guided by the cognitive principle of relevance (following the path of least resistance and looking for the most significant cognitive effects), the inferences involved in comprehension are split into non-sequential stages or 'sub-tasks' in which the recipient (the audience or hearer) constructs and compares "…anticipatory hypotheses about the overall structure of the utterance being processed" (Curcó, 1995, p. 31). These include: a "hypothesis on explicit content (explanations) via decoding, disambiguation, reference resolution, and other pragmatic enrichment processes"; a "hypothesis on the intended contextual assumptions (the premises involved)"; and a "hypothesis on the intended contextual implications (the implied conclusions)" (Wilson and Sperber, 2004, p. 615). Communication is achieved when the inferential process leads the hearer to recover the message that the speaker wanted to convey.

A relevance-theoretic account of linguistic creative humour

One of the first attempts to apply the framework of RT to an account of humour in communication was undertaken by Maria Jodłowiec, who tried to "…characterise pragmatic mechanisms (...) involved in the production and comprehension of verbal jokes" (1991, p. 242). As shown by Francisco Yus' (2017) recent overview, many other studies have followed this relevance-theoretic orientation, both pursuing a similar goal or exploring kindred subjects such as: (a) the perception of something as humorous (whether intended or not); (b) possible classifications of humour; (c) kinds of constraints and effects involved in humorous comprehension; (d) humorous ironies and narratives; (e) conversational humour; (f) the translation of humour. We will not dwell on many aspects of these interesting subjects. Part of what is expected from RT as "a pragmatic theory of verbal humour" is that it "should be able to predict what kind of utterances and texts will be humorous and [explain] why" (Curcó, 1997, p. 165).

We will confine our attention to a relevance-theoretic explanation of intentional linguistic creative humour.

A special feature of such an explanation — unlike some other linguistic approaches to the subject — is that it takes a decisively cognitive shift. It claims that "rather than assuming that being humorous is a property of texts, and hence concentrating on their structure, (…) what we need to understand to characterize verbal humour are the mental processes [involved] when humorous effects are derived" (Curcó, 1997: 165). As we stressed in our characterisation of 'creative humour', the intentional production of humorous effects (insofar as it is a creative process) should bear the hallmark features of creativity: it must be novel, have some value, and spark surprise. As we argued, some of those features are achieved by creative humour (stemming from surprise triggered by the incongruity-resolution account; value whether as cognitive debugging or as emotional relief). Accounting for those features in a systematic understanding that is in keeping with the natural sciences would amount to naturalising creativity.

RT provides a useful framework for systematically unifying what the production of novel instances of linguistic humorous incongruities requires. In cases of intentional production of linguistic humour "the hearer has been forced [by the speaker…] to interpret the utterance responsible for the humorous climax as consistent with the principle of relevance, [in a way that] contradicts some other assumption either explicitly conveyed by an immediate utterance, or manifest in the accessible context of interpretation" (Curcó, 1997, p. 30). Thus a surprising incongruity is produced by design, with the initial information presented by the speaker. The comedian's interaction with the audience triggers his listeners to make specific anticipatory hypotheses, in which he exploits the cognitive relevance mechanism. These anticipatory hypotheses, created relying on the mind-reading ability of the speaker to set the hearers' interpretation on a specific cognitive path, generalise expectations concerning the type of information to follow; "however, people often encounter new information (…) that deviates from expectations" (Wyer and Collins, 1992, p. 665). Thus, in linguistic creative humour the speaker sets up a resolvable incongruity by violating expectations of relevance, intentionally triggering the audience to revise their

ideas in the light of a new piece of information. Here, the speaker exploits his ability to guess what hearers are thinking in order to 'reverse-engineer' their inferential strategies and how they assimilate information. Since the inferences involved in comprehension require several non-sequential stages or 'sub-tasks', humorous incongruities may arise due to several non-equivalent factors, such as: the resolution of lexical and syntactical ambiguities; determinations of reference assignment; disambiguation; free enrichment and conceptual adjustment. This process can be structurally described in terms of lexical and syntactic calculations. Nevertheless, the main cognitive mechanism is usually tweaked through contextual information for the inferential strategies that the speaker believes are most likely to be used by his listeners (as we saw in the examples of I-RTs).[8] However, in order to achieve the surprise required for creative humour, at least some aspects of the inference should be unanticipated by the audience; otherwise the whole thing would not be funny. Thus, RT shows how (notwithstanding the relevance-based algorithmic processing of information) creative humour makes people laugh by spawning novel, unpredictable psychologically incongruities whose resolution turns out to be (by design) cognitively valuable for the intended audience.

We have seen that RT promises to uncover the cognitive mechanisms used by speakers to intentionally produce creative instances of linguistic humour. We assumed that a theory explaining this kind of creativity might hold out hope of a fully-fledged cognitive explanation of overall creativity. Nevertheless, it seems that linguistic creative humour as portrayed by RT can be explained as a cognitive phenomenon. Unfortunately, it turns out that linguistic humour is highly dependent on the speaker's ability to steer the hearer's inferential strategies and his access to context along predictable paths to 'solve' an incongruity. This makes it unlikely that such a theory of creativity would have any predictive value. On the positive side, it does reveal a naturalistic mechanism for

humour, embodying the unpredictability of outcomes and explaining their cognitive value. Thus, the apparent paradox for a theory of creativity vanishes.

CONCLUSIONS
In this paper we explored some aspects of the scientific study of creativity by focusing on examples of linguistic humour. We argued that this sort of creativity can be accounted for within the framework of a valuable cognitive approach but that does not yield a recipe for producing novel humorous situations. We first identified three great puzzles that arise in attempting to naturalise creativity. After having explained each of them, we outlined some promising ways in which they have been addressed by cognitive studies of creativity. In a second section, we focused our attention on the links between creativity and humour. Then we restricted our discussion to what we called 'creative humour'. In the final section, we showed how an important cognitive approach to human communication can account for creative humour.

If Relevance Theory does explain linguistic creative humour, being able to intentionally produce humorous linguistic episodes could depend on being able to systematically elicit amusing incongruities. However, the information required for creating some kinds of linguistic humour might only be available to someone who takes part in and is in the midst of conversational contexts. If this is indeed the case, it would be impossible for the information to be anticipated or generalised in the way that Sheldon Cooper's Unified Theory of Comedy seems to require "…to elicit laughter from anyone at any time". These remarks should give us some pause for thought on what (not) to expect from a naturalisation of creativity. On the one hand, such an explanation may not entail predictive or implementation capabilities. On the other hand, the fact that creativity appears within a systematic theory involving physical entities and mechanisms might be an indication that its explanatory power relies precisely on something that precludes being able to forecast its outcomes.

---

8   For the reconstruction of several case studies using the theoretical framework of RT, see Curcó (1995; 1997) and Yus (2016; 2017).

BIBLIOGRAPHIC REFERENCES

Amabile, T. (1987). The Motivation to be Creative. In S. Isaksen (ed.), *Frontiers of Creativity Research: Beyond the Basics* (pp. 223-254). Buffalo: Bearly Limited.

Amir, O., and Biederman, I. (2016). The Neural Correlates of Humor Creativity. *Front. Hum. Neurosci*, *10*, 597.

Arieti, S. (1976). *Creativity: The Magic Synthesis*. New York: Basic Books.

Beermann, U., and Ruch, W. (2009). How Virtuous is Humor? Evidence from Everyday Behavior. *Humor: International Journal of Humor Research*, *22*(4), 395-417.

Bermúdez, J. (2014). *Cognitive Science. An Introduction to the Science of Mind.* (2nd ed.). Cambridge: Cambridge University Press.

Bermúdez, J. (2005). *Philosophy of Psychology. A Contemporary Introduction*. New York: Routledge.

Boden, M. (2004). *The Creative Mind: Myths and Mechanisms*. London: Routledge.

Boden, M. (2009). Creativity: How Does It Work? In M. Krausz, D. Dutton and K. Bardsley (eds.), *The Idea of Creativity* (pp. 237-250). Leiden: Brill.

Boden, M. (2016). *AI: Its Nature and Future*. Oxford: Oxford University Press.

Carroll, N. (2014). *Humour: A Very Short Introduction*. Oxford: Oxford University Press.

Chan, Y-C., Chen, H-C., and Lavallee, J. (2013). The Impact of Gelotophobia, Gelotophilia and Katagelasticism on Creativity. *Humor: International Journal of Humor Research*, *26*(4), 609-628.

Cropley, A. (2011). Definitions of Creativity. In M. Runco and S. Pritzker (eds.), *Encyclopedia of Creativity* (2nd ed.) (pp. 511-524). San Diego: Academic Press.

Curcó, C. (1995). Some Observations on the Pragmatics of Humorous Interpretations: A Relevance-Theoretic Approach. *UCL Working Papers in Linguistics*, *7*, 27-47.

Curcó, C. (1997). *The Pragmatics of Humorous Interpretations: A Relevance-Theoretic Account*. Unpublished PhD Dissertation. London: University College London.

Davies, G., and Subkoviak, M. (1975). Multidimensional Analysis of a Personality-Based Test of Creative Potential. *Journal of Educational Measurement*, *12*(1), 37-43.

Dietrich, A., and Kanso, R. (2010). A Review of EEG, ERP, and Neuroimaging Studies of Creativity and Insight. *Psychological Bulletin*, *136*, 822-848.

Estany, A., and Herrera, R. (2016). *Innovación en el saber teórico y práctico*. London: College Publications.

Getzels, J., and Jackson, P. (1962). *Creativity and Intelligence*. New York: Wiley.

Gibson, J. (2019). *An Introduction to the Psychology of Humor*. New York: Routledge.

Gordon, W. (1962). *Synectics*. New York: Harper.

Hobbes, T. (1651). *Leviathan*. Cambridge: Cambridge University Press, 1991.

Hong, N. (2010). Mow'em All Down Grandma: The "Weapon" of Humor in Two Danish World War II Occupation Scrapbooks. *Humor: International Journal of Humor Research*, *23*(1), 27-64.

Howard-Jones, P., Blakemore, S., Samuel, E., Summers, I., and Claxton, G. (2005). Semantic Divergence and Creative Story Generation: An fMRI Investigation. *Cogn. Brain Res*., *25*, 240-250.

Hurley, M., Dennett, D., and Adams, R. (2011). *Inside Jokes: Using Humor to Reverse-Engineer the Mind*. Cambridge: MIT Press.

Jodłowiec, M. (1991). What Makes Jokes Tick? *UCL Working Papers in Linguistics, 3*,241-253.

Kellner, R., and Benedek, M. (2016). The Role of Creative Potential and Intelligence for Humor Production. *Psychology of Aesthetics, Creativity and the Arts*. Advance online publication.

Koestler, A. (1964). *The Act of Creation*. London: Hutchinson.

Kronfeldner, M. (2009). Creativity Naturalized. *The Philosophical Quarterly*, *237*(59), 577-592.

Limb, C., and Braun, A. (2008). Neural Substrates of Spontaneous Musical Performance: An fMRI Study of Jazz Improvisation. *PLoS ONE*, *2*(3), e1679.

Mankoff, R. (2009). Foreword. In J. Morreall, *Comic Relief: A Comprehensive Philosophy of Humor* (pp. ix-x). Malden: Wiley-Blackwell.

Maslow, A. (1971). *The Father Reaches of Human Nature*. New York: Viking Press.

Morreall, J. (2009). *Comic Relief: A Comprehensive Philosophy of Humor*. Malden: Wiley-Blackwell.

Murdock, M., and Ganim, R. (1993). Creativity and Humor: Integration and Incongruity. *Journal of Creative Behavior*, *1*(27), 57-70.

O'Quin, K., and Derks, P. (2011) In M. Runco and S. Pritzker (eds.), *Encyclopedia of Creativity* (2nd ed.) (p. 511-524). San Diego: Academic Press.

Papineau, D. (2006). Naturalist Theories of Meaning. In E. Lepore and B. Smith (eds.),*The Oxford Handbook of Philosophy of Language* (pp. 175-188). Oxford: Clarendon Press.

Rouff, L. (1975). Creativity and Sense of Humor. *Psychological Reports*, *37*, 1022.

Scruton, R. (1982). Laughter. *Proceedings of the Aristotelian Society*, *56*, 197-212.

Sperber, D., and Wilson, D. (1987). Précis of *Relevance: Communication and Cognition. Behavioral and Brain Sciences, 10*, 697-754.

Sperber, D., and Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed.). Oxford: Blackwell.

Sprevak, M. (2016). Philosophy of the Psychological and Cognitive Sciences. In P. Humphreys (ed.),*The Oxford Handbook for the Philosophy of Science* (pp. 92-114). New York: Oxford University Press.

Stokes, D. (2007). Incubated Cognition and Creativity. *Journal of Consciousness Studies*, *14*, 83-100.

Stokes, D., and Paul, E. (2016). Naturalistic Approaches to Creativity. In J. Sytsma and W. Buckwalter (eds.), *A Companion to Experimental Philosophy* (pp. 318- 333). Malden: Wiley-Blackwell.

Torrance, P. (1966). *Torrance Tests of Creative Thinking*. Princeton: Personnel Press.

Treadwell, Y. (1970). Humor and Creativity. *Psychological Reports*, *26*, 55-58.

Ward, E. H., Hickey, K. J., and Richards, K. T. (2020). *Drug Pricing and Pharmaceutical Patenting Practices*. Washington: Congressional Research Service.

Wilson, D., and Sperber, D. (2004). Relevance Theory. In L. Horn and G. Ward (eds.), *The Handbook of Pragmatics* (pp. 607-632). Oxford: Blackwell.

Wyer, R., and Collins, J. (1992). A Theory of Humor Elicitation. *Psychological Review*, *99*(4), 663-688.

Yus, F. (2016). *Humour and Relevance*. Amsterdam: John Benjamins.

Yus, F. (2017). Relevance-Theoretic Treatments of Humor. In S. Attardo (ed.), *The Routledge Handbook of Language and Humor* (pp. 189-203). New York: Routledge.

Ziv, A. (1976). Facilitating Effects of Humor on Creativity. *Journal of Educational Psychology*, *68*(3), 318-322.

BIOGRAPHICAL NOTE

Mario Gensollen

Has a PhD in Philosophy (Universidad Nacional Autónoma de México) and is a Tenured Lecturer in the Philosophy Department at Universidad Autónoma de Aguascalientes (since 2006). He is Visiting Lecturer at Universitat Autònoma de Barcelona. His research focuses on Social Epistemology, particularly on the intersections between Epistemology, Argumentation Theory and Political Philosophy. For further information, see https://mariogensollen.academia.edu

Marc Jiménez-Rolland

Has a PhD in Humanities (Universidad Autónoma Metropolitana). He is Assistant Lecturer in the Philosophy Department at Universidad Autónoma de Aguascalientes (since 2008). He was Adjunct Lecturer at Universidad Autónoma de Zacatecas (2011-2012). His research focuses on Epistemology and the Philosophy of Science, especially on formal approaches to problems in these areas. For further information, see https://uaa.academia.edu/MarcJiménez