Synthese (2010) 175:193–218 DOI 10.1007/s11229-009-9498-8

Darwinian 'blind' hypothesis formation revisited

Maria E. Kronfeldner

Received: 26 August 2008 / Accepted: 3 March 2009 / Published online: 11 April 2009 © The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract Over the last four decades arguments for and against the claim that creative hypothesis formation is based on Darwinian 'blind' variation have been put forward. This paper offers a new and systematic route through this long-lasting debate. It distinguishes between undirected, random, and unjustified variation, to prevent widespread confusions regarding the meaning of undirected variation. These misunderstandings concern Lamarckism, equiprobability, developmental constraints, and creative hypothesis formation. The paper then introduces and develops the standard critique that creative hypothesis formation is guided rather than blind, integrating developments from contemporary research on creativity. On that basis, I discuss three compatibility arguments that have been used to answer the critique. These arguments do not deny guided variation but insist that an important analogy exists nonetheless. These compatibility arguments all fail, even though they do so for different reasons: trivialisation, conceptual confusion, and lack of evidence respectively. Revisiting the debate in this manner not only allows us to see where exactly a 'Darwinian' account of creative hypothesis formation goes wrong, but also to see that the debate is not about factual issues, but about the interpretation of these factual issues in Darwinian terms.

Keywords Darwinism \cdot Blind variation \cdot Creativity \cdot Hypothesis formation \cdot Guided variation \cdot Lamarckism \cdot Evolutionary epistemology \cdot Popper \cdot Campbell \cdot Simonton

Popper (1972, 1974a,b, 1975, 1984) and Campbell (1960, 1974a,b) both claimed that creative hypothesis formation shares a fundamental feature with biological evolution: it is based on 'blind' variation. This analogy has caused a lot of confusion and critique.

M. E. Kronfeldner (🖂)

Max Planck Institute for the History of Science, Berlin, Germany e-mail: mkronfeldner@mpiwg-berlin.mpg.de

What does 'blind' mean and, considering its definition, is that the way we produce new ideas in science? And if we do produce new ideas blindly, does the analogy between hypothesis formation and biological evolution add something to our knowledge of how new hypotheses arise in the minds of scientists?—These are the two main questions this paper aims to address with respect to the blind-variation analogy, the analogy between creative hypothesis formation and novelty in biological evolution.

I am only concerned with this particular analogy. Other analogies between science (or culture) and evolutionary biology might well be valuable even if the blind-variation analogy is ill-founded. Whether these other analogies make sense cannot be addressed here. Notable analogies are for instance Hull's (1988, 2001) evolutionary model of science, focusing on the evolutionary concept of fitness, to illuminate competition and cooperation in science, or models of cultural evolution, using quantitative population-based models to study gene-culture interaction and to illuminate similarities and differences in the inheritance patterns of culture and biological evolution.¹ Both these evolutionary approaches are not only independent of the blind-variation analogy; representatives of these models are themselves among the critics of this specific analogy.² Finally, this paper looks exclusively at scientific creativity. Many of the points that will be mentioned could certainly also be made for art or other cultural domains, but to show this has to wait for another occasion.

In Part 1, I shall propose three distinctive meanings of the concept of blind variation, i.e. three ways how the blind-variation analogy can be made precise. I shall then introduce and develop the critics' point of view in Part 2, which is that creative hypothesis formation is guided rather than blind. In Part 3, I shall discuss and assess three compatibility arguments. These say that the critique of guided variation can be made compatible with the view that there exists an important analogy in terms of the blindness of variation between creative hypothesis formation and biological evolution.

By revisiting the debate in this manner, I aim to provide a new and systematic discussion of the arguments used for and against the claim that creative hypothesis formation is blind. But by systematically discussing the issue in the described manner, I also aim to introduce a new meta-level to the long-lasting back and forth in the debate, a level at which we can address what *follows* from the critique of guided variation for the 'blind variation' analogy with respect to hypothesis formation. Towards this end, the paper integrates contemporary developments in philosophy and psychology of creativity, a body of literature by and large ignored by philosophers of science. Integrating these developments allows us to see that important matters of fact are actually clear and accepted by both sides; only the interpretation varies. To a considerable degree the debate is only about whether the use of the concept of blind variation, taken from evolutionary biology, is justified or not, given the factual basis accepted by both sides

¹ Advocated by Boyd and Richerson (1985), Richerson and Boyd (2005), and others (e.g. Mesoudi et al. 2006).

² See Hull (2001, pp. 513–514) for critique of blind variation and matters of independence; Boyd and Richerson (1985, pp. 81–98) already distanced themselves from 'blind' variation; see Mesoudi (2008) for the most recent account of the independence of cultural evolution from the blind variation analogy. If I understand him correctly, he defends this via (what I call here) 'selectionist compatibility' and 'bias compatibility.' He is more focused on cultural evolution and concentrates thus on the outcome of cognitive processes (and their optimality) and less on how creative outcomes are cognitively produced.

of the debate. Distinguishing clearly between different meanings of blind variation will prove to be crucial in order to untangle the factual from the conceptual issue. If these distinctions, instead of the metaphorical term 'blind' were used, it would have been recognized easily that the debate is characterised by 'much ado about almost nothing', i.e. about a claim few would dare to put into question.

1 Three meanings of 'blind' and the critique of guided variation

In Campbell's terminology, to which Popper has referred as well, biological evolution and creative hypothesis formation are both instances of a so-called "blind-variation-selective-retention-process." When applied to creative hypothesis formation, 'variation' stands for the process of generating novelty, meaning the modification or creation of hypotheses. 'Selective retention' stands for cognitive intentional selection of the generated items: after generating hypotheses, we select them, we decide to keep or to discard them. The meaning of the term 'blind' is less clear. A complete historical review of Popper's and Campbell's usage of the term would by far exceed the space available here. My aim is not historical exegesis; I shall follow a more systematic road.

1.1 The concept of blind variation in evolutionary biology

To evaluate an analogy the base of the analogy must be clear and defined independently of the analogy itself. Otherwise the definition of the concept that is building the heart of the analogy, and is supposed to apply to the base and the target of the analogy, is in danger of being manufactured so that the analogy works, i.e. in a questionbegging manner. The only approach that prevents a question-begging definition of blind variation is to use the biologists own concept of blindness, and to check on this basis whether the target of the analogy (creative hypothesis formation) matches onto the biological base.

In evolutionary biology, as it developed from Darwin until today, the term 'blind variation' refers to the claim that new variants occur in an *undirected* manner. A new variant occurs undirected, if it occurs randomly with respect to its adaptivity. As the biologist Futuyma (1998) puts it, variation or "mutation is random in the sense that the chance that a particular mutation will occur is not influenced by whether or not the organism is in an environment in which that mutation would be advantageous" (Futuyma 1998, p. 282; emph. in the orig.). In formal terms, as philosopher Sober (1992, p. 39) puts it: "Let u be the probability of mutating from A to a and v be the probability of mutating from a to A. Mutation is directed if (i) u > v and (ii) u > vbecause w(a) > w(A), where w(X) is the fitness of X." If that is not the case, the mutation is undirected. In such a case variation and selection are "decoupled," as Toulmin (1972, p. 338) has put it. Decoupling means that the two "twin sub-processes of variation and selection [...] take place *quite independently*, so that the factors responsible for the selective perpetuation of variants are entirely unrelated to those responsible for the original generation of those same variants" (Toulmin 1972, p. 337). If the chance of the occurrence of a new variant is *not influenced* by the factors that determine the selection of the variant, then variation and selection are decoupled and the change

undirected. Given this definition, directedness is a matter of degree, whereas undirectedness is simply the absence or negation of any directedness and thus (logically) not a matter of degree.

Three further points about this concept of undirectedness have to be taken into account to prevent widespread misunderstandings surrounding the idea of undirected and directed variation. First, directed changes are often called 'Lamarckian.' Lamarck's concept of directedness, however, was stronger. Lamarck assumed that new characteristics arising in an organism are directed in the sense that the factors causing new characteristics determine the course of evolution without any selection process that weeds out non-adaptive features. With respect to adaptive evolution, Lamarck assumed a factor that automatically produces adaptive features in organisms, a factor that *induces* adaptive features. In terms of the concept of coupling, Lamarck's concept of directedness would mean that the coupling is 100 per cent: that the productive factors *are* the selective factors, that u = 1 and u = 1 *because* of w(a) > w(A). A system with total coupling would make selection superfluous. But as long as coupling of productive and selective factors is not 100 per cent, the change can still be a Darwinian selection process, even though the process would not rely on undirected changes as defined above. This shows that a *partial* coupling is compatible with Darwinism in a wide sense, i.e. with any theory that assumes that a sorting process (selection or drift) is a necessary part of biological evolution. Undirectedness is thus not a necessary ingredient of Darwinism in that wide sense.³ Yet, undirected variation is still fundamental to contemporary Darwinism (what I shall call Darwinism in the narrow sense), i.e. the Darwinism you find in a standard textbook version of contemporary evolutionary biology, such as in Futuyma (1998). And the reason why it is fundamental to contemporary Darwinism in this narrow sense (i.e. Modern Synthesis Darwinism) is simply: there is broad consensus that undirectedness holds for biological evolution.⁴

Second, undirectedness does not mean *randomness* in the sense that all mutations or variants are *equally likely*. Although mutation is stochastic, there is no equiprobability of specific mutations, even though they occur in an undirected manner. Different loci or regions of loci in a genome have different mutation rates. Environmental factors, such as ultraviolet radiation, chemical mutagens, or nutrition can increase the mutation rate at a certain loci. There is also the phenomenon of mutation pressure where the forward mutation rate at a locus is different from the backward mutation rate. Consequently, undirected variation has to be distinguished from *random variation* in

³ See Kronfeldner (2007, pp. 499–500) for a more detailed account of the compatibility of directed variation with Darwinism in this wide sense. According to this account, Darwin was a Darwinist in this wide sense since he partly relied on directed variation.

⁴ See Sniegowski and Lenski (1995) for a detailed account of this Modern Synthesis consensus, still accommodating seemingly 'directed/adaptive mutations' (e.g. due to environmental stress) by alternative explanations along the lines of the consensus. It should nonetheless be mentioned that the Modern Synthesis view is currently under severe attack. As I understand it, this attack from ecological and developmental perspectives targets the gene selectionism inscribed in it, regarding the gene as the sole locus of evolutionary change and heredity. Even if a move away from gene centrism would include directed *phenotypic* variation as a factor in biological evolution, this does not challenge the proposition that mutations are undirected, and it would certainly not change the *concept* of undirectedness. According to this, the modern synthesis would have to be extended but not changed.

the sense of equiprobability of trials. To stress this is important since the two have often been confused. A statement from Campbell illustrates this point. To say that blind variation means that "specific correct trials are no more likely to occur at any one point in a series of trials than another, nor than specific incorrect trials" (Campbell 1960, p. 93) is misleading, since it suggests that undirected variation requires that specific adaptive trials are as likely as specific maladaptive ones, i.e. that they are equiprobable.⁵ Recall, mutation is undirected if it is *not* the case that (i) u > v and (ii) u > v because w(a) > w(A), where w(X) is the fitness of X. This definition states that if there is a higher probability of u, it must be excluded that this is *because* of a higher fitness of the allele 'a' compared to the allele 'A'. It does not follow that we have directed variation, if u > v. If a specific adaptive trial is more likely than any other trial this is not necessarily a case of directed variation. Therefore, it is not the case that undirected variation requires that specific adaptive trials are no more likely to occur than specific maladaptive ones. In summary: no equiprobability is required for undirected variation, and if there is a higher probability of some variants, it is not necessarily due to coupling, i.e. due to the influence of the selective environment of this variant.

Last but not least, directed variation has to be distinguished from developmental constraints, which can also lead to a higher probability of some variants. At a certain point of time in evolution, certain variants are not possible or at least very unlikely. The reason for this is that if the respective developmental pathways were missing, an occurring mutation would not lead to a viable variant. A common example is the effect of a half-wing, which prevents the future appearance of an arm in place of a wing. The half-wing acts as a developmental constraint. Thus, it is not the case that at each point in time every conceivable phenotypic variation can occur for selection to operate on. This limit on the range of variation, a statistical bias in variation, which is exhibited by biological evolution, is compatible with Darwinism-even in its narrow sense, in sharp contrast to coupling. Developmental constraints are part of standard evolutionary biology, even if they diminish the explanatory force of the selective part of the overall evolutionary process.⁶ Conceptually, developmental constraints can be distinguished from coupling in the following way: first of all, developmental constraints do not influence the occurrence of mutations or re-combinations. In principle, they could, but as far as I know, they are not proven to do so. They merely influence whether a mutation can be expressed and whether the change, if expressed, will have a positive effect on the fitness of the organism. Second, and more important here, developmental constraints are (if at all) the effects of *past* selective environments. Coupling is different in both respects: first, coupling would make adaptive variation more likely to occur in the first place, i.e. at the genotypic level. Second, coupling would not be an effect of *previous* selection. On the contrary, it would be a direct effect of the *current* selective environment. This difference is often overlooked, since both-developmental

⁵ A similar misleading formulation occurs, for instance, in Mesoudi (2008, pp. 244, 252), despite the other merits of that paper.

 $^{^{6}}$ See Dawkins (1986, pp. 307–312), Amundson (1989), or Futuyma (1998, p. 276) on this particular issue.

constraints and directed variation due to coupling—lead to a statistical bias in variation. To remember the difference will prove to be decisive in Part 2.

1.2 Application to creative hypothesis formation

If we now transfer all this to creative hypothesis formation, then the following holds: creative hypothesis formation is blind, in the sense of an *undirected* change in ideas, if the change is decoupled, i.e. if the occurrence of new ideas is *not* influenced by factors that determine the selection of these new ideas.

Again, variants do not have to be produced at random in order to be undirected. This is not required for biological evolution either. In principle, we are certainly free to define blind variation differently, e.g. as requiring *random variation*; but it is important to distinguish such an interpretation of 'blind variation' clearly from undirected variation.⁷ And it is even more important to note that we would then ignore the base of the putative analogy.

Whereas randomness is a stronger requirement than undirectedness, there is also a third interpretation of blind variation that does not even demand undirectedness. As a reaction to complaints that his concept of blind variation offered earlier is unclear, Campbell (1974b) stated that for him blind means "*unjustified*": hypotheses are *trials*, and that means that they are *unjustified by origin*, i.e. they are not directly caused or warranted by their origin in observation. Blind-as-unjustified variation means that we do not know *in advance* whether the ideas we produce are worthwhile. This does not exclude a partial coupling, since a partial coupling can still lead to variants that are not worthwhile. Popper also referred to this concept of blind-as-unjustified variation. A main goal of his evolutionary epistemology was to argue against what he called observationism or the bucket theory of the mind, which assumes that the ideas we come up with are—in principle—warranted by their origin in observational sense data. For Popper, this idea is tantamount to Lamarck's idea of an environment *inducing* adaptive changes in organisms without the help of selection.

How, and if at all, beliefs can get their epistemological warrant, if not by their origin, is an epistemological question that is irrelevant for this study.⁸ Consequently, whether the term 'unjustified' is useful or whether we should, given the problem of induction and related problems, get rid of such terms is irrelevant here as well. The important thing for this study is that we clearly see the difference between *unjustified*, *undirected*, and *random* variation. Unjustified variation refers to the fact that we come up with ideas concerning which we do not know in advance whether they are worthwhile.

If we try to find an analogue to this concept of blind-as-unjustified variation in the biological realm, it would amount to a denial of a Lamarckian 100% coupling. We

⁷ Kim (2001, pp. 109, 113), for instance, uses blind variation as random variation vis-à-vis and undistinguished from other meanings. Others, however, did acknowledge the distinction: Popper did (e.g. 1974a, p. 117; 1974b, p. 35), and so did Blute (1979, pp. 44–45), and Campbell in various places (Campbell 1960, p. 92; 1974a, p. 56; 1974b, p. 148), despite the confusing remark mentioned above.

⁸ Campbell, for instance, argued for a more constructivist warrant of knowledge. See the papers in Heyes and Hull (2001).

would then have to count everything as 'blind' that is not caused in a Lamarckian inductive manner. Cases of partial coupling, i.e. cases of directed variation, are then still cases of 'blind variation,' and blindness becomes a matter of degree. This stands in sharp contrast to the biologist's definition of blindness as undirectedness, which restricts 'blind variation' to the *absence* of *any* coupling. This contrast points to an important difference in perspective. The two interpretations of the qualifier 'blind', undirected and unjustified, relate to two opposed partitions of the continuum from zero to 100% coupling. The concept of undirected variation treats zero coupling as the lonely member of one class and puts everything else in a second class of direct-edness; the concept of unjustified variation puts 100% coupling in an isolation camp and everything else into the crowded blindness camp.⁹

To make this change of perspective explicit, and to decide for one partitioning, is of utmost importance in order to find a straight way out of the back-and-forth of the debate and to prevent falling back into the metaphorical language of being 'blind'. Since it is equally important to use a non-question begging point of view, the biologist's concept of blindness is, all things considered, the safer route.

Let me briefly outline the consequences of this terminological decision. If we concentrate on creative hypothesis formation, the concept of unjustified variation counts all cases of hypothesis formation as 'blind', in which we do not know in advance whether the trial-ideas will prove worthwhile. The concept of undirected variation, the one used by biologists, would however exclude most of these cases as being 'blind'. As illustrated at the beginning of this Sect. 1.2, if we transfer the concept of undirected variation, then creative hypothesis formation is blind only if the change is decoupled, i.e., if the occurrence of new ideas is *not* influenced by factors that determine the selection of these new ideas. If there is coupling between producing and selecting factors in the formation of an idea, the generation of the idea is not undirected and therefore not 'blind' (given the biologist's definition of blind variation), even though the creator might not know in advance whether the idea is worthwhile, i.e., even though the generation of the idea might well be a case of unjustified variation.

2 Guided variation

The proposal that creative hypothesis formation is blind (in the biologist's meaning of the term) has been objected to repeatedly. According to critics, creative hypothesis formation is directed, or *guided*, as most prefer to say in the case of cognition. It is directed since variants that are useful are more likely to occur *because they are thought to be useful as a solution to a given problem.*¹⁰ What critics have usually left unexplained is *how* exactly the orientation towards a problem leads to an adaptive bias, and whether *all* cases of creative hypothesis formation include such an orientation

⁹ I want to thank one of the anonymous referees of this paper particularly. By proposing an alternative definition of undirectedness, he helped me, precisely because of his divergent perspective, to see the importance of this point.

¹⁰ See Ruse (1974), Richards (1977), Skagestad (1978), Thagard (1988), Amundson (1989), Sternberg (1998), and Dasgupta (2004), to name but a few. A similar critique, but not focused on the cognitive level, can be found in Boyd and Richerson (1985, pp. 81–98).

towards a problem. In the following, I shall summarize the critique and the evidence used for it, and develop the critics' point of view in the direction of an answer to the latter two questions.

2.1 Guided variation in creative hypothesis formation

The critique of guided variation in hypothesis formation rests on two kinds of evidence: first, historical cases as paradigm examples of creativity, and, second, more direct and general evidence from cognitive psychology, which systematically studies the cognitive processes involved in hypothesis formation and other creative cognition. I shall argue that the two kinds of evidence support each other.

Before I do this, a third kind of evidence needs to be mentioned. It is used by those critics that focus on diffusion of ideas and analyze cultural change as a population-level evolutionary process. Boyd and Richerson (1985, pp. 81–98), for instance, have claimed that cognition is guided by various criteria (i.e., cognitive and other biases) and that therefore the outcome of a cognitive process is usually directed. What is then fed into the populational process can be already biased towards adaptivity and can thus increase the probability that adaptive items spread at the populational level via social learning. As I have argued elsewhere in more detail (Kronfeldner 2007, pp. 507–09), this critique does not address guided variation directly at the cognitive level and is thus compatible with the claim that the cognitive process reaches the 'guided variation' that is fed into the population by a truly undirected variation and selection process at the cognitive level. Adaptivity at the end of the cognitive process will thus not provide evidence that helps to solve the debate about creativity of individual people, a debate that can only be solved by focusing on the cognitive level and how individuals arrive at the ideas that are then fed into the public domain of shared knowledge.¹¹

The claim of those critics that focus on the cognitive level is that through the orientation towards a certain problem, certain knowledge gets activated, and this knowledge structures the search space for solutions and, finally, triggers the *production* of certain ideas and not others. Completely strange hypotheses are not generated at all; on the contrary, usually 'good guesses' are generated. My claim is that the guidedness that is apparent in these good guesses is due to *cognitive coupling*: part of the selective environment (already acquired knowledge) influences the production of the new thought trials.

2.2 A historical illustration

Take one of the many well-known cases from history of science as an illustration, e.g. Kekulé's discovery of the structure of the benzene molecule. Given our historical evidence of Kekulé's career and what we know about how he arrived at the hypothesis for the structure of the benzene molecule, it is clear that Kekulé's general knowledge about chemistry and his knowledge that benzene has this and that chemical features

¹¹ See Mesoudi (2008) who further discusses the precise nature of guided variation at the level of cultural evolution and how it even directs biological evolution via niche construction.

(derived from earlier rounds of hypothesis formation and selection) played a decisive role in the production of the hypothesis that benzene forms a ring. The discovery was not instantaneous and not independent of the knowledge he acquired before.¹² On the contrary, the knowledge enabled some ideas and not others. After producing the hypothesis, it was evaluated. And most important here, it was done so with the help of the same cluster of knowledge, which helped produce the new hypothesis, namely, knowledge about chemistry and chemical features. Because of this cognitive coupling between generating and selecting factors, i.e., because of the influence of already acquired knowledge, the hypothesis about the structure of the benzene molecule has to be counted as being generated in a directed manner.

Nonetheless, the influence of already acquired knowledge in such cases is still "plastic," as Briskman (1981, p. 148), defender of the Darwinian account of creativity, suggests, since the already acquired pieces of knowledge "may themselves have to be modified" to reach a viable hypothesis. In addition, and quite important for this study: other factors, factors that are not yet part of previously acquired knowledge, also play a part in the occurrence (as well as the selection) of novel ideas.

If we take Kekulé's case to represent the standard case of creative problem solving, then three points have to be acknowledged: that guidedness (but not blindness) is a matter of degree, that hypotheses are not generated in an undirected manner, and that the already acquired knowledge alone does not explain what happens in the mind of creative problem solvers. After illustrating these three points in a row, I will justify why I believe that Kekulé can count as an example for how creativity usually works, even though we do not know what really happened in his mind.

Partial coupling in hypothesis formation shows that guidedness is a matter of degree, i.e. there can be more or less directedness. But note that the same does not hold for undirectedness: there *cannot* be more or less blindness *in the sense of undirectedness*, since undirectedness is defined as the absence of *any* coupling. This is often ignored (e.g. when defenders of blind hypothesis formation want to save the blind variation analogy by reference to 'degrees of blindness') (e.g. Simonton 1999a, p. 62; 2005, pp. 300–301). Either a process is undirected, or it is more or less directed. What defenders of the blind variation analogy say (i.e., that there are 'degrees of blindness') is wrong, but what they *want* to say (i.e., that constraints do not equal total coupling or foresight) is correct, as I shall concede in more detail in Part 3.

If formation and selection of a hypothesis is coupled, and if formation and selection of biological variation is not coupled at all, which I take for granted here, then there exists an important disanalogy to the case of biological evolution.

Finally, those factors involved in creative cognition that are not part of already acquired knowledge can be regarded as coincidental. They are coincidental in the sense that they were, prior to the hypothesis formation, not part of the knowledge space defining the problem the scientist tries to solve (e.g. for Kekulé's case, to find a description for the structure of the benzene molecule). It is these factors, and the

¹² See Findlay (1968, pp. 34–41) for a short survey on the historical case. See Schaffer (1994) for a lucid critical discussion of the 'anecdotes' surrounding Kekulé's dreams and imaginations, and for further references for the historical case. See Kronfeldner (forthcoming) on a 'demystified' description of Kekulé's imaginations from the point of view of contemporary psychology of creativity.

cognitive mechanism related to them, that are today often cited as explanation for the distinctive spontaneity and surprise so often accompanying creative cognition. And, important for this study, this spontaneity and surprise exists despite guidedness. Last but not least, the explanation via coincidental factors is taken into account by critics as well as defenders of blind variation. The interpretation in Darwinian terms is controversial: the Darwinian interprets the influence of these coincidental factors as evidence for 'blind variation' and the non-Darwinian denies this. This already indicates that much of the debate, as stated in the introduction, is not, at least not today, about the available evidence in Darwinian terms. But before this can be defended, we have to place the argument from guided variation on a more stable foundation.

2.3 Towards a more general account of guided variation

We do not know exactly what happened in Kekulé's mind, and one or even a few historical examples might not suffice to make a convincing argument for guided variation, as Simonton (2005), defending blind variation, stressed against Dasgupta (2004), who offered three further historical cases along the line of the argument from guided variation. Yet a short review of the evidence coming from detailed comparative historical studies, from experimental cognitive science, and from computer studies, supports the assumption that Kekulé's example represents a paradigm case of creative cognition.

The case-study based historical approach to creativity that compares paradigm individual examples of creativity, occurring in the wilderness of real science, art, etc., is despite its limitations still regarded as an ample supplier of data for understanding creativity, especially since it is the only approach for which it is beyond doubt that it studies truly creative people. This historical approach is probably best exemplified by the work of Gruber (Gruber 1974; Gruber and Davies 1988; Gruber and Wallace 1999), and Gardner (1993). By careful comparative studies, on Darwin and others, Gruber established that creativity in individuals gradually evolves over time and that this involves the "loose coupling of the evolving sub-systems of knowledge, purpose, and affect" (Gruber and Davies 1988, p. 265). Gardner's extensive case studies also confirm the importance of expert knowledge in creativity. For instance, it usually takes a decade of 'aspiration,' of studying the accumulated and received knowledge of the domain, before a 'great man of history' can celebrate his 'coming out' with an important 'inspiration.' All this evidence from historical case studies supports the stance of guided variation.

Another supplier of evidence is cognitive science, which tries to understand creativity 'in the lab,' in part with the help of computer models. Cognitive approaches to creativity disagree on many detailed points.¹³ Yet, there seem to be two points of convergence: firstly, in cases of problem solving, previously acquired knowledge structures the search space so that 'good guesses' are more likely to occur, and, secondly, many cognitive mechanisms are involved in this.

¹³ See Runco (2007, pp. 1–38), Ward et al. (1999), or Sternberg (1988, 1998).

a considerable amount of knowledge of a specific domain) usually show a higher performance level than novices if they are presented with problems that demand a more or less creative solution.¹⁴ And even defenders of blind variation admit that expertise plays a structuring role in creative problem solving, as I will illustrate in Part 3.

Research on particular cognitive mechanisms points in the same direction. The cognitive mechanisms currently debated with respect to creative cognition are numerous: perception, visual imagery, associational linkage, conceptual combination, analogical reasoning, abstraction, use of metaphors, conceptual expansion, memory retrieval, spreading activation, defocused attention, reduced latent inhibition, and opportunistic assimilation, to name but a few. These multiple mechanisms are interacting in a complex and highly efficient manner and at the moment of insight, the sudden 'Eureka,' at a quite high speed, some of them on auto-pilot, i.e. without conscious monitoring. They are held to explain not only creativity in the lab, but also the myriads of idiosyncratic and unpredictable cases of real creativity in the wilderness of human endeavours, cases that are considered as hard to bring under a general model, precisely because of the sheer complexity of interaction of these mechanisms.¹⁵

None of these mechanisms has been proven to work in an undirected manner. They all build on mental structures, i.e., already acquired knowledge, leading to a partial coupling as described above for Kekulé's case. For most of these cognitive mechanisms (e.g. analogical reasoning) the influence of already acquired knowledge has never been much in dispute; but some have been used explicitly for claiming that creative cognition is 'blind.' They have thus taken centre-stage within the debate. Simonton (2005, pp. 302, 305, 306), for instance, refers to spreading activation, defocused attention, and reduced latent inhibition as "blind," while at the same time acknowledging that this does not exclude the possibility that already acquired knowledge has a structuring influence on these mechanisms. But as Schooler and Dougal (1999, p. 352) note, spreading activation is "by no means random in nature. Rather, the direction and extent of the spread of activation critically depends on (a) the specific items that were initially activated and (b) the underlying structure of an individual's knowledge representation." Seifert et al. (1995) show the same for the mechanism of opportunistic assimilation, supporting thereby that 'chance favours only the prepared mind.' As Schooler and Dougal (1999, pp. 352–353) summarize: "just because creativity can be fostered by random cues does not necessarily implicate randomness in the psychological process of creativity."¹⁶ To give a full account of all the experiments and mechanisms currently discussed in creative cognition approaches, would by far exceed the space available here. But what can be concluded from the short review of the state of the art is: there is a convergence in cognitive studies of creativity towards

¹⁴ See Sternberg (1998, p. 163) for an overview of studies on expert-novice differences, in reply to the Darwinian model of undirected variation.

¹⁵ See Sternberg (2006, p. 433), Boden (2004, pp. 260–276), and Ward et al. (1999, pp. 190–191); see also Mumford (1999), Seifert et al. (1995), and Weisberg (1993). The accounts of these authors certainly differ but these differences can safely be ignored here.

¹⁶ See Mumford (1999, p. 345) for the same argument with respect to associational linking and conceptual combination.

a multiple mechanisms view; most of these mechanisms also occur in normal cognition; their complex interaction helps to explain the characteristic surprise in creative cognition; and, finally, even the most characteristic mechanisms of creative cognition, such as opportunistic assimilation, show the guidedness postulated by critics of the blind variation claim.

Research on computer models of creativity also points in the same direction. First, even a generalist 'creative computer program' (i.e. EURISKO), not relying so much on domain-specific expert knowledge, is equally working in a constrained manner. "[C]onsiderable specialist knowledge has to be provided if it [the creative computer program, MK] is not to generate hosts of nonsensical (as opposed to merely boring) ideas" (Boden 1998, p. 351).¹⁷ In addition, when we try to find a solution for a certain problem, it seems that we are not working like genetic algorithms, which are programs explicitly designed as Darwin machines, instantiating the undirected variation-selection process found so effective in biological evolution, as Boden (1999, pp. 366–369, 2004, pp. 318–322) summarizes this domain of research.¹⁸ Genetic algorithms are "blatantly Darwinian," as Simonton (1999b, p. 313) states correctly; yet we seem not to be like genetic algorithms. If we were, this would be evidence for undirected variation in cognition that would have strength comparable to the classical experiments conducted in biology to validate the view that mutations are undirected.¹⁹ But even Simonton (1999b, p. 313), currently the hardest defender of the blind-variation claim within psychology of creativity, admits that other non-Darwinian computer programs can equally model our creativity. Boden, defender of a connectionist model of creativity, concludes from her account of computer creativity (be it classical, connectionist, or evolutionary), that our creativity is based on generative structures that constrain the production of ideas so that certain ideas are more likely to occur than others, even though our mind often uses coincidental clues, as do many computer programs. Note that this conclusion does not entail that genetic algorithms cannot be creative, or that they are useless for understanding our creativity. On the contrary, they help us to see differences between genetic algorithms and our minds; and this in turn helps us to construct genetic algorithms that more closely model our creativity. For instance, those genetic algorithms that are considerably constrained by expert knowledge (e.g. by an "aesthetically structured space," as is the genetic algorithm used by the artist William Latham for a painting program) seem to do just that: they produce less radical variants, but those that are produced can be the basis for incremental and systematic development (Boden 2004, p. 320). And this is something we usually do with our most creative ideas.

Even Simonton admits that guided variation is generally exhibited in cases of problem solving (e.g. Simonton 2003, pp. 316–317, 2005, pp. 300–301). This is a further indication that it is not controversial anymore whether cases like Kekulé should be described the way critics do, or whether we have reliable evidence for guided

 $^{^{17}}$ In his critique of evolutionary epistemology, Thagard (1988) has claimed the same for his computer model of creativity.

¹⁸ A classical work on genetic algorithms is: Holland (1975); Compare Koza (1992) and Fogel (1999).

¹⁹ See Luria and Delbrück (1943). For a review of the debates surrounding these experiments see Sniegowski and Lenski (1995).

variation from history and cognitive science. What follows for the blind-variation analogy (and the corresponding claim that creativity is analogous to biological evolution) is controversial, and also, as part of this issue, whether there is any additional evidence that allows us to make the evidence for guided variation irrelevant for the analogy. These issues will be dealt with in Part 3. But before we can move on to this part, a specific complication must be addressed.

There is one issue that still stands in the way of a general account of guided variation. That the orientation towards a certain problem makes creative hypothesis formation guided does not entail that *all* cases of creative hypothesis formation include such an orientation. There might well be cases of true serendipity, cases of creative hypothesis formation that are *defined* as not having an orientation towards a specific problem, cases where you find something without even looking for it as a solution to a certain problem (Roberts 1989). Alexander Fleming's serendipitous discovery of penicillin is a standard example of true serendipity in science. If one is not looking for a solution to a specific problem, then the critique of guided variation, at least in the form introduced above, cannot be applied in the same sense as it applies to Kekulé, a clear case of problem-oriented hypothesis formation.

What does serendipity show? Not much, since cases of serendipitous hypothesis formation are very likely also directed. If serendipitous findings can count at all as hypothesis *formations* (instead of counting them as discoveries without prior hypothesis formation), then the occurrence of the hypothesis at the moment of finding (e.g. Fleming's idea that this thing over there is a useful thing for this and that) is still influenced by prior knowledge acquisition. And the prior knowledge might influence the creative process in a way that makes the *recognition* of something *as a useful discovery* (in our case the thing then called penicillin), and thus the formation of the hypothesis, more likely. This only requires that serendipitous *recognition* counts as a part of hypothesis *formation* and not as a part of hypothesis *selection*. What serendipity then shows is that guided hypothesis formation is likely even without a *prior* orientation towards a certain problem.

To summarize: in creative hypothesis formation (in standard problem-oriented cases as well as in cases of serendipitous hypothesis formation, if there are any), the occurrence of hypotheses is directed. *If*, as is still believed to be the case and taken for granted here, biological evolution is based on undirected variation, then there is an important difference between creative hypothesis formation and biological evolution, a difference that destroys the adequacy of the analogy between the two phenomena.²⁰

3 Three compatibility arguments

The main answer to the critique of guided variation has always been that there is, nonetheless, an important and adequate analogy. Defenders of the blind variation claim do not deny that hypothesis formation is guided by previously acquired knowledge, but

 $^{^{20}}$ There might well be other differences between the base and the target of the analogy that would not destroy the adequacy of the analogy. That creative hypothesis formation does not rely on sexual reproduction is, for instance, irrelevant for the adequacy of the analogy (or claim) that creative hypothesis formation is as undirected as biological variation.

they deny that this destroys the adequacy of the analogy between novelty in human minds and novelty in biological evolution. The defensive claim is: it does not follow from the apparent guidedness that creative hypothesis formation and biological evolution are not analogous. Popper (1974b, p. 117 or 1975, pp. 84–87) already admitted in 1974 that creativity is guided in the way critics claim, even though he sticks to the claim that creativity is based on blind variation (and therefore similar to biological evolution). The question that will be important for the rest of the paper is thus: what follows from guided variation for the claim that creativity is similar to the Darwinian process of undirected variation and consequent selection? It will prove to be decisive that we clearly distinguished undirectedness from other meanings of blind variation, and from developmental constraints.

To approach the issue in a systematic manner, I shall highlight three arguments that state that guided variation is compatible with the thesis that creativity is analogous to biological evolution. The first argument states (what I call) a selectionist compatibility, the second a bias compatibility, and the third a hidden chaos compatibility.

3.1 Selectionist compatibility

The *selectionist compatibility* argument states that a bias within variation is compatible with the Darwinian view of evolution as a *selective process*. Popper and Campbell go in this direction. All that is required for a process to be a Darwinian evolutionary one—and therefore parallel to biological evolution—is (a) the exclusion of foresight in a narrow sense and (b) the exclusion of a Lamarckian-like process. These have to be excluded since both would guarantee that adaptive new features occur without the help of selection, i.e. without weeding out errors. From this point of view, the only thing that is required is that variation is unjustified in Campbell's sense, unjustified by origin. Unjustified variation necessitates selection of a hypothesis, and thus justifies the analogy between creativity and biological evolution.

3.1.1 The exclusion of foresight in a narrow sense

Popper said that for any given problem one cannot "anticipate its (unknown) solution otherwise than by a fortunate accident" (Popper 1974a, p. 36). Campbell said: "Real gains must have been the products of explorations going beyond the limits of *foresight* or *prescience*, and in this sense blind" (Campbell 1960, p. 92; emph. added). Foresight, in the narrow sense as Popper and Campbell assume it here, does not refer to the 'good guesses' that are part of guided variation. It refers to real *prescience*.²¹ If we take this into account, we can see that Popper and Campbell merely offer a *conceptual* argument about creativity: real gains *must* have been 'blind,' by definition. If creativity is *defined* as going beyond what is already known, that is, defined as bringing about something *new* ('real gains'), then creativity excludes foresight, since foresight is defined as knowing something already. As Plato stressed in his paradox of search

²¹ Mesoudi (2008) uses the term clairvoyance for prescience, stressing the importance of distinguishing between clairvoyance on the one hand, and good guesses (what he calls foresight) on the other.

in the *Meno*: if we know something, we cannot look for it since we know it already; if we, however, do *not* know it, we cannot even look for it since we do not know what we are looking for. That is why we cannot bring about something new by will and foresight. The most curious thing about this specific aspect of the concept of creativity, assumed by Popper and Campbell and quite widespread in philosophy of creativity, is that a perfect God-like creator can thus not be creative, since a perfect being would have foresight and could thus prevent errors. For a God-like creator there is no need to be creative, so to say.²²

But why should foresight in the narrow sense be so important?—Because, if we had it, we could prevent errors, just as a God-like creator. Thus, the lack of foresight in finding a creative solution means that the moment we come up with an idea, we do not know whether the idea is worthwhile. This necessitates cognitive selection of thought trials and is the basis for why Campbell calls the origin of new variants unjustified. Creativity thus—by necessity—involves a selection process. Creativity does not however require by definition that we produce ideas in an *undirected* manner. Campbell's argument is supported not only by common sense, holding that we just make errors; it is also supported by empirical research on bounded rationality and biased decision making. We simply are not perfect: we make errors and are biased in many ways.²³

But note that selection of hypotheses is a serial rather than a parallel mode of selection. Nature operates with parallel processing: many organisms function as simultaneous trials and the best variants increase in frequency due to selection. Human minds are usually not engaged in such a kind of parallel processing. We do not breed ideas in our minds that then increase their frequency so that many copies of an idea float around in our minds. Even though we sometimes compare different ideas that seem to be candidate solutions for one and the same given task, we usually develop alternatives diachronically, one after another. Creativity is—as Nickles (2003, p. 63) says—an instance of "serial evolution." We have an idea and try it out. If it fails, we consider another idea at hand, maybe even a previously rejected one, or we search for a new alternative that is not yet at hand, and thus try another yet unwarranted idea, until we reach a stable resting place. This procedure is evidently a selection process.

3.1.2 Exclusion of observationism

Popper is most prominent for stressing a further aspect of human creativity, creativity as opposed to what he called observationism or psychological induction, which he judges to be parallel to Lamarckism. As mentioned above, Lamarck explained local adaptation through pointing to the moulding influence of the environment. Through this influence an organism would be able to directly adapt to its environment. As Lamarck conceived evolution, selection was not necessary for evolution, since each organism was thought to interact with its environment directly and to reach 'harmony' through this interaction. According to Darwin, adaptive and non-adaptive features of organism arise in a population and the good ones have to be selected, if evolution should

²² For more on the underlying concept of creativity see Kronfeldner (forthcoming).

²³ Mesoudi (2008, pp. 249–252) summarizes the respective empirical research on rationality, discovery, and decision-making relevant for this point.

happen. After repeated selection over generations, this indirect 'trial-and-error' leads to adaptations. Popper states an analogous contrast: hypothesis formation cannot be reduced to gathering "'direct knowledge' of anything immediately 'given'"; we are not "passive recipients of information impressed upon us from outside" (Popper 1974b, p. 116). Instead, we *create* ideas and select them afterwards. The hypotheses we build on these ideas are—in Campbell's words—'unjustified,' i.e. not guaranteed by origin to be adaptive for the problems that we want to solve with them. Popper concludes: "[W]e must abandon any approach which starts from sense data and the given, and replace it by the assumption that all human knowledge is fallible and conjectural. It is a product of the method of trial and error" (*ibid*.) Popper therefore stresses an old contrast: learning by experience, i.e. learning by passively copying the environment, is not creative. If creativity is defined as excluding copying from the environment, then it holds that, if hypothesis formation is creative, it excludes—by definition—observationism, a kind of psychological Lamarckism.²⁴

What do (a), the exclusion of foresight (in the narrow sense), and (b), the exclusion of psychological induction, show for the questions at issue here? They illustrate that human hypothesis formation is *neither* analogous to a hypothetical God-like creation of hypotheses, with perfect knowledge and true foresight, *nor* analogous to a Lamarckian direct adaptation to environment. However, this does not prove that it is Darwinian *in the narrow sense*: Darwinian as biological evolution has turned out to be, Darwinian as contemporary Darwinism in biology defines it, i.e. as requiring undirected variation. Creative hypothesis formation is an instance of a coupled selective process and biological evolution is not. An evolutionary account of hypothesis formation can thus not be called evolutionary in the same narrow sense as contemporary Darwinism is called evolutionary. In other words, as long as biological evolution rests on undirected variation, there is no strong analogy between creative hypothesis formation and biological evolution.

But there is not only Darwinism in the narrow sense. As explicated in Part 1, Darwinism in the wide sense does not necessarily require undirectedness, even if biological evolution is still believed to fulfil this rather strong criterion. It merely requires that selection is not superfluous. By pointing to the fact that creativity is—as biological evolution—necessarily a selection process, we can thus restore the claim that there is an analogy and reach a certain compatibility between the claim that creative hypothesis formation is blind and the claim that it is guided: it is guided, but based on blind, i.e. unjustified variation.²⁵

Yet, even if we achieve compatibility of this sort, the price for it is a certain trivialization of the claim. First of all, in no way does the claim that creativity is a selective process *explain* creativity: it states that creativity excludes foresight and psychological induction. However, this is not yet an explanation of the *cognitive processes* involved in creativity. It merely excludes two kinds of explanation: (a) an explanation through

²⁴ Popper was not the first who stated this contrast: The argument goes back to James (1880) who accused Herbert Spencer of being a psychological Lamarckist.

 $^{^{25}}$ This is in line with the consequences of the perspective that Hull et al. (2001) have defended by claiming that it is not a *direct analogy* that is decisive for the comparison between creativity and biological evolution, but a *general* account of selection processes.

foresight, similar to explaining biological adaptations in a creationist way; (b) an explanation through a reduction to direct information transmission from the environment, similar to explaining biological adaptations in a Lamarckian way. Second, the claim that ideas can fail is much weaker than the claim that they are generated in an undirected manner. Even though the critics' claim (that creative hypothesis formation is guided) does not contain a full explanation either, it nonetheless says more about creative hypothesis formation than the claim that it involves a selection process—and so would the (false) claim that it is undirected. Compared to the latter, the claim that creative hypothesis formation is based on unjustified variation has much less explanatory force. Moving back to unjustified variation in face of the critique of guided variation is thus answering a critique by trivializing the original claim. You may secure the analogy with this strategy, but you lose most of the explanatory force of the analogy.²⁶

3.2 Bias compatibility

The second defence against the critique of guided variation, the bias compatibility argument, states that the difference between the alleged decoupled selection processes in biological evolution and the coupling in creativity is an illusion. According to this, biological evolution *also* exhibits a bias that makes adaptive features more likely. What happens in creativity is thought to be the same as what happens in biological evolution due to the effect of cumulative selection. Guidedness in creativity should thus be explained as the perfectly Darwinian effect of cumulative previous variation-selection cycles. Already Campbell (1974b, p. 151) reacted to the critique of guided variation in this way by pointing to his nested hierarchy of vicarious selection processes. He claimed that there is a cascade of ontogenetic selection processes (behavioural trial-and-error due to locomotion, internalised trial-and-error due to sensory systems, trial-and-error at the level of 'thought trials'). According to him, this nesting, i.e. previous selection processes at the same or at vicarious levels, leads to the guided variation critics described; but at least one of the vicarious processes was itself blind. The products of previous selection stages then restrict further 'thought trials' in the way critics claimed for creative hypothesis formation. This can be interpreted as trying to restore the analogy by showing that there is a blind variation process *some*where in the hierarchy. Yet this cannot restore the claim that there is blindness at the *level of hypothesis formation* and only this has been questioned by the argument from guided variation.

Stein and Lipton (1989) further developed Campbell's strategy of defence. They try to neutralise what they call the "anomaly of guided variation" by stating that biological evolution is not "truly blind" either, since it is constrained in various ways

²⁶ Despite the explanatory weakness of the claim that creativity is a selection process, it might well have had a *heuristic* value within the psychology of creativity. Whether it has had such a heuristic value cannot be decided philosophically. Evidence for it could only be provided by detailed empirical case studies on whether the 'discovery' of specific mechanisms (e.g. reduced latent inhibition or spreading activation, postulated as helping to explain creativity) were encouraged by the idea that creativity is based on a Darwinian selection process.

(Stein and Lipton 1989, p. 43). They are not alone in claiming this. As mentioned, Simonton (1999a, p. 62) claims that despite guided variation, there are "degrees of blindness," in biological evolution as well as in creativity. He combines this with the thesis that the degree of blindness is dependent on how much the variation is restricted by constraints that are the effect of previous trials.²⁷

Stein and Lipton (1989) mainly refer to pre-adaptations; one of the examples they use is the half-wing, acting as developmental constraint in biological evolution, enabling and constraining the subsequent evolution of the wing. Pre-adaptations are taken to be analogous to previously acquired knowledge in hypothesis formation. In both cases, the developmental constraints operative due to the pre-adaptations can explain how the variation that is available for selection at a certain step of the overall cumulative process can be biased towards certain variants. On this basis, Stein and Lipton want to restore the claim that there is a strong similarity between creativity and biological evolution, a similarity that entails more than the claim that both are selective processes. According to them, both exhibit a bias that can be explained in the same way—through the effect of developmental constraints operative in biological or epistemic pre-adaptations. Guided variation can then be explained (and thus accommodated) as a standard Darwinian effect of cumulativity.

Their argument rests on a simple but central error. They ignore that developmental constraints have to be distinguished from coupling: the former is an effect of *previous* cycles of variation and selection and the latter an effect of the current selective environment of a respective type of entity whose evolution is at issue. The bias compatibility argument says that-despite guided variation-creativity is analogous to biological evolution, since in both cases there is a bias in variation that can be explained as the standard effect of previous cycles of variation and selection. On the one hand, this is true; on the other hand, it is not. The question is whether creativity is, first, an instance of coupling (and therefore not similar to biological evolution), or whether it is, second, an instance of a cultural analogue of developmental constraints. Now, the problem is that creativity is both at the same time. The knowledge, i.e. the selective environment, that causes coupling between variation and selection, is the knowledge that is the effect of previous cognitive selection. These effects are 'pre-adaptations' that restrict future developments. They certainly do this through a cultural analogue to developmental constraints: knowledge and cognitive heuristics, as well as certain standards of a domain (epistemic virtues like testability in science, aesthetic virtues in the arts, for instance).²⁸ Now, the essential difference between creativity and biological evolution is that in creativity the *present* selective environment contains the effects of *previous selection*. This is not necessarily the case for biological evolution. The tree in the selective environment of a giraffe is not an effect of previous selection rounds of

²⁷ Compare Simonton (1988, pp. 4–5, 1995, pp. 473–474, 1999b, p. 311, 2003, pp. 316–317, 2005, pp. 300–301), (Nickles, 2003, pp. 65–66), Buskes (1998, pp. 115–123), or Simon (1981, pp. 205–206), who also used a 'selectionist' analogy, but without relying on, or explicitly denying undirected variation. Mesoudi (2008) has put forward the so-far most extended and up-to-date version of such a bias compatibility argument, mixed with a selectionist compatibility argument. He addresses the impact of guided variation for population-level cultural selection processes and forms thus a complement to this paper, which focuses on the individual level of creativity.

²⁸ For more on the analogous effects of 'developmental constraints' in culture see Wimsatt (1999).

the giraffe as a type. Yet, our previously acquired knowledge (comparable to the tree) is indeed the *effect of previous selections* of the body of belief, whose evolution is at issue (analogous to the giraffe). Owing to this difference, developmental constraints and coupled evolution are hard to distinguish in creativity and easy to distinguish in biological evolution.

Since creativity contains developmental constraints and coupling at the same time, critics and defenders are both right. The bias compatibility argument is however still ill founded. Stein and Lipton state that coupling is no argument against a strong similarity, since the bias in creativity can be explained as being analogous to developmental constraints. Therefore, they state, "we do not need the coupling of variation and selection" to explain guided variation (Stein and Lipton 1989, p. 53). We do not need coupling to explain the *presence* of the knowledge that acts as current selective environment; this is explained by cumulativity. But we need coupling in order to explain how the knowledge that is the effect of previous selection leads to guided variation: how it causes the bias in variation. Stein and Lipton overlook this precisely because they fail to differentiate consistently between developmental constraints and coupling. This is the central error in the bias compatibility argument.

In conclusion, Stein and Lipton mistakenly explain a factual instance of coupling as merely analogous to the effects of cumulativity in biological evolution. They ignore the fact that biological evolution is an instance of *decoupled cumulative evolution*, while creativity is a special case of *coupled cumulative evolution*. The bias compatibility argument fails.

3.3 Hidden chaos compatibility

There is one chance left to answer the critique of guided variation. The *hidden chaos compatibility* argument states that the apparent guidedness of creativity is based on an unconscious, i.e. hidden chaos. This argument has been defended most prominently by the creativity psychologist Dean K. Simonton in his "chance-configuration theory" (1988, 1995) and goes back to Poincaré's (1908) theory of unconscious idea formation.²⁹

Neither Poincaré nor Simonton deny that the previously acquired knowledge leads to coupling, i.e. that it can constrain the production of thought trials in the way we described it above for Kekulé. What they deny instead is that this contradicts the thesis that there is an *unconscious* process of truly 'blind' variation as part of the overall process that is guided by selective criteria, which represent already acquired knowledge. Previously acquired knowledge might, for instance, influence which elements of ideas enter the unconscious process in order to get reshuffled and newly combined, while the combinations themselves are formed blindly, as Poincaré (1908, pp. 393–394) puts it.

It is unclear whether Poincaré and Simonton demand that the unconsciously produced "chance-configurations" are really produced at random or in an undirected

²⁹ The argument is implicit already in Campbell (1960, 1974a,b). I will use the term "hidden chaos" from Stein and Lipton (1989, p. 38), who also defend such a compatibility. Nonetheless, in this case, I will concentrate on Simonton's more detailed version.

manner. What Simonton, for instance, explicitly demands is that a "large number of potential permutations exist, all with comparably low but nonzero probability" (Simonton 1988, p. 7). Be it as it may, we can assume that the unconscious selection process is based on many variants, or on variants that are generated in an undirected manner, or even in a random manner. In all these cases, the problem is the same: the empirical evidence for such a claim about an *unconsciously* produced *waste* of ideas, ideas that never come to consciousness, is hard to get.

Poincaré himself admitted that he could not cite evidence for his account; he admitted that it is only a hypothesis. Campbell (1960, 108f) conceded that his model of creativity as blind variation is still not an explanatory psychological theory, since it would have to prove and specify the exact cognitive mechanisms involved in creativity. But this was in the 1960s and psychology of creativity has come a long way since then. It is Simonton who claimed that in light of the current developments in psychology of creativity Poincaré's model turns out to be the correct explanation of creativity. Other creativity theorists resist this claim and state that there is no unambiguous evidence for a special process of unconscious trial-and-error.

According to the cognitive approaches to creativity mentioned in Part 2, creativity can be reduced to the complex interaction of a group of cognitive processes, most of which are operative in all kinds of reasoning, not only in creative reasoning. At the same time there is abundant evidence that creativity involves what psychologists call instantaneous insight: ideas occur so suddenly that it feels as if they are not produced at will. The reports of such sudden insights, from Archimedes 'Eureka' to Kekulé, Poincaré, Einstein, or Watson and Crick, still fill contemporary books on creativity, be they historically or psychologically minded. Insight can thus count as a phenomenological datum of creativity research. It usually occurs after two loosely defined stages of the overall creative process, a preparation stage and a break called incubation stage. The special process of unconscious variation-and-selection is thought to provide an explanation for insight and for what happens during incubation. Yet, in addition to the unconscious variation-and-selection hypothesis there are other common hypotheses about insight, as Seifert et al. (1995) stress. One of the alternatives is the opportunistic-assimilation hypothesis. It states that it is not incubation itself and a hypothetical unconscious hidden chaos, but the assimilation of serendipitous, i.e. coincidental clues that explains sudden insight phenomena. By a series of careful experiments, Seifert et al. (1995) show that insight is not more likely to occur just by extending or shortening the incubation stage, which is the supposed time when the hidden chaos is thought to form itself. On the contrary, they showed that insight is only then considerably more likely to occur, when an impasse is reached during preparation and when the individual has then the chance to be influenced by coincidental clues during incubation. The increased probability, which occured when an impasse was reached and coincidental clues were available, was not predicted and cannot be explained by the unconscious variation-and-selection hypothesis. In other words, the predictions derived from an unconscious variation-and-selection hypothesis were not confirmed, whereas the predictions of the opportunistic-assimilation hypothesis were. Even though this does not prove that no hidden chaos occurs, it does support the claim that the process of opportunistic assimilation (of accommodating coincidental new information to a given structured problem space) is essential. And, as mentioned in Part 2, opportunistic assimilation itself is a mechanism far from being random or undirected.

The evidence for the hidden chaos model that Simonton (1999a,b) offers from his own research is historiometric. The problem is that although this evidence is indeed consistent with the hidden chaos theory, as Simonton (1999b, p. 316) has claimed, it is *also consistent with alternative approaches* that do not rely on a hidden chaos. Let me concentrate on the most striking issue. Simonton has cited his own historiometric evidence for the "equal-odds rule," which shows, for instance, that the expected probability of career success remains constant within a career, regardless of the creator's age and increase in accumulated knowledge. In other words, the proportion of hits (defined quantitatively via citation statistics) stays the same, even if more previously acquired knowledge constrains the psychological make-up of a person.

Whatever the value of the historiometric data is, the data cannot provide evidence for a special subconscious process of variation-and-selection. The historiometric evidence is too *indirect* to be conclusive in that respect. First of all, it measures how often we succeed according to quantitative criteria. It does not explain how a person comes up with those ideas that this person judges himself as worthwhile, which may well include the successful and the unsuccessful ones. Second, the random pattern is compatible with the alternative explanation of insight mentioned above. It might well be that the proportion of Kekulé's successful ideas did not increase over his career and that Kekulé produced his idea of the benzene ring the way the alternative non-evolutionary model suggests, namely with the help of opportunistic assimilation and normal cognitive processes such as memory, perception, abstraction etc., working in a highly efficient manner and at the moment of insight at quite a high speed, and some of them, like perception, on auto-pilot (i.e. not consciously monitored). Third, the historiometric model aggregates data over a variety of contexts, a procedure that might well lead to a random pattern merely because of the aggregation. The complex interaction of cognitive processes and social processes in these contexts alone lets us "expect the resulting data to fit a random model," irrespective of any assumptions about the cognitive processes involved, as the psychologist Mumford (1999, p. 344) argued. Thus, the historiometric evidence is consistent with and predictable from models that do not assume that there is a special process of unconscious trial-and-error.³⁰

Finally, it should be noted that Simonton (e.g. 1999b, p. 312) also refers to experimental evidence to defend his Darwinian account of creativity, for instance, to opportunistic assimilation. Yet this has been formulated as an explicit alternative to the chance-configuration mechanism. Opportunistic-assimilation can only support unjustified variation, i.e. selectionist compatibility, but not a hidden chaos.

As long as there is no conclusive experimental evidence for an unconscious production of a 'hidden chaos,' selectionist compatibility is the only resting place for creativity as a Darwinian process.

 $^{^{30}}$ For a critique of non-aggregative historiometric evidence provided by Simonton (2007a), derived from a single case, see Mumford and Antes (2007).

3.4 Moving back to selectionist compatibility

This argument is supported by a more sweeping look on Simonton's work over the years, since it seems that it is this weaker analogy to which Simonton actually committed himself, an analogy to which he explicitly moved back to in his later papers, withdrawing from the claim that creativity is based on the chance-configuration mechanism. The chance-configuration mechanism is then treated as but one of many mechanisms that could account for the selective process (Simonton 1999a,b, 2003, 2007b). In Simonton's own words, the stress of 'blindness' of variation in creativity only denotes "the lack of foresight in the production of variations-the inability to generate purposively the most adaptive variations. [...] The term blindness also has the advantage of not committing the theory to any particular variation mechanism" (Simonton 1999b, p. 310; emph. added). The evidence offered in the same paper, and discussed above, would then have to be interpreted as evidence for the weaker claim that creativity is a selection process. Yet, nobody disagrees with this claim. In the context of contemporary psychological explanations of creativity, this weaker claim becomes superfluous, precisely because, as Simonton himself says, it is *compatible* with many of the diverse mechanisms that are suggested by psychology of creativity as playing a role in creative achievements and which are not in any sense 'blind' or 'evolutionary.' This compatibility with diverse cognitive mechanisms further corroborates the claim that the selectionist claim itself does not contribute much to an explanation of creativity (see above 3.3).

In 2005 then, Simonton (2005, p. 307) decided to "drop the explicit BVSR model and its Darwinian theoretical foundations," and to explain "creativity in terms of 'combinatorial models that treat the phenomenon as a 'constrained stochastic process'." The systematic analysis offered here makes clear *why* he *had* to do so: because of the compatibility of unjustified variation with diverse cognitive mechanisms, it becomes superfluous to call the mechanisms themselves 'blind' or 'Darwinian' despite their myriad differences at the cognitive level.³¹ Last but not least, his final withdrawal illustrates that unjustified variation is all we can safely postulate for creativity. Unfortunately, the latter is only implicitly admitted by Simonton, who goes on to use the term 'blind' in all three meanings of random, undirected, and unjustified variation, and does so interchangeable, without clearly distinguishing between them.

4 Conclusion

Let me collate the results: There are three quite distinct meanings of blind variation, i.e. unjustified, undirected, and random variation. The critique of guided variation attacks undirected variation and claims that guidedness shows a disanalogy between biological evolution and hypothesis formation. There are three strategies of defence, three compatibility arguments. The *selectionist compatibility* argument states that a bias within variation is compatible with the Darwinian view of evolution as a

³¹ Simonton (2005) gives ideological reasons for his withdrawal: given that Darwinism is for many an ideologically suspicious thing, it is wise not to use the language.

selective process; the *bias compatibility* argument states that the guidedness in creativity is analogous to developmental constraints in biological evolution; the *hidden chaos compatibility* argument states that the apparent guidedness of creative hypothesis formation is based on a hidden, unconscious blind variation process.

Each compatibility argument has considerable problems: the selectionist compatibility claim holds but trivializes the original claim, the bias compatibility claim rests on the simple but central error of confusing coupling of variation and selection with developmental constraints, and the hidden chaos compatibility claim lacks evidence from psychology of creativity.

Since selectionist compatibility is the only compatibility that holds, let me end by summarizing the trivialization inherent in it. The kind of blind variation that can be defended (i.e. unjustified variation) is compatible with guided variation and with diverse cognitive processes. Two basic consequences derive from each of these two compatibilities, one about the descriptive and one about the explanatory force of the analogy between creative hypothesis formation and biological evolution. Because of the first compatibility (i.e. selectionist compatibility), the analogy has some *descriptive* adequacy, but not very much, since the critics' point still holds: creative hypothesis formation exhibits coupling, while biological evolution does not, given the current consensu in biology. Because of the second compatibility (i.e. compatibility with diverse cognitive processes), moving back to the thesis that creativity involves good guesses (i.e. unjustified leaps into the unknown followed by cognitive selection) leads to the loss of almost all the explanatory force of the analogy. Although unjustified variation excludes certain explanations (i.e. a God-like formation of ideas with foresight and an observationist account of idea formation), it does not itself offer an explanation of creativity, which would have to be given at the level of the diverse cognitive processes. It is the compatibility or non-commitment with respect to cognitive mechanisms that actually destroys almost all of the explanatory force of the claim that should be defended.

Blind variation as unjustified variation transferred back to biological evolution simply means that the process is a selection process. This meaning is part of any theory that can rightfully be called Darwinian and is fulfilled by creative hypothesis formation as well as biological evolution. But since creative hypothesis formation is not based on undirected variation, one has to be careful not to assume more than that. Since selectionist compatibility is a rather weak claim, and since all parties to the debate actually take guided variation for granted, large parts of the debate should be regarded as 'much ado about nothing'. We can choose to call creativity a Darwinian selection process, but if we refrain from doing so, we do not lose anything from our understanding of creativity, at least nothing that we did not know already or independently of any Darwinian framework, namely that we are simply not perfect and make lots of errors.

Acknowledgements I have presented earlier versions of this paper at the ISHPSSB 2005 Meeting at the University of Guelph and at the Philosophy Department at the University of Regensburg. I want to thank all participants of these meetings for their helpful comments. I am highly indebted to Hans Rott for his severe criticism and propulsive help, whenever it was necessary for my work. I am also obliged to Marion Blute, Sören Hägquist, Christophe Heintz, Mohammadreza Memarsadeghi, and Francesca Merlin, and to the anonymous referees: they all helped a lot to develop my arguments and to make them more precise. An earlier version of the paper has been awarded the "Karl Popper Essay Prize 2006–2007".

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Amundson, R. A. (1989). The trials and tribulations of selectionist explanations. In K. Hahlweg & C. Hooker (Eds.), *Issues in evolutionary epistemology* (pp. 413–432). Albany: SUNY Press.
- Blute, M. (1979). Learning theory and the evolutionary analogy. *Cogprints* (1979), 2002. http://cogprints.ecs.soton.ac.uk.
- Boden, M. A. (1998). Creativity and artificial intelligence. Artificial Intelligence, 103, 347-356.
- Boden, M. A. (1999). Computer models of creativity. In R. J. Sternberg (Ed.), Handbook of creativity (pp. 351–391). Cambridge: Cambridge University Press.
- Boden, M. A. (2004). *The creative mind: Myths and mechanism* (2nd ed.). London and New York: Routledge.
- Boyd, R., & Richerson, P. J. (1985). Culture and the evolutionary process. Chicago: University of Chicago Press.
- Briskman, L. (1981). Creative product and creative process in science and art. In D. Dutton & M. Krausz (Eds.), *The concept of creativity in science and art* (pp. 129–155). Dordrecht: Martinus Nijhoff Publ.
- Buskes, C. (1998). The genealogy of knowledge: A Darwinian approach to epistemology and philosophy of science. Tilburg: Tilburg University Press.
- Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, 67, 380–400. (Reprinted in and quoted from *Evolutionary epistemol*ogy, rationality, and the sociology of knowledge, by G. Radnitzky & W. W. Bartley, 1987 (pp. 91–114) LaSalle: Open Court).
- Campbell, D. T. (1974a). Evolutionary epistemology. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (Vol. 1, pp. 413–463). La Salle: Open Court Publ.
- Campbell, D. T. (1974b). Unjustified variation and selective retention in scientific discovery. In F. Ayala & T. Dobszhansky (Eds.), *Studies in the philosophy of biology: Reduction and related problems* (pp. 139–161). London: Macmillan.
- Dasgupta, S. (2004). Is creativity a Darwinian process? Creativity Research Journal, 16, 403-413.
- Dawkins, R. (1986). The blind watchmaker. London: Penguin.
- Findlay, A. (1968). A hundred years of chemistry (3rd rev. ed.). London: Duckworth.
- Fogel, L. J. (1999). Intelligence through simulated evolution: Forty years of evolutionary programming. New York: Wiley.
- Futuyma, D. J. (1998). Evolutionary biology (3rd ed.). Sunderland, MA: Sinauer Associates.
- Gardner, H. (1993). Creating minds: An anatomy of creativity seen through the lives of Freud, Einstein, Picasso, Stravinsky, Eliot, Graham, and Gandhi. New York: Basic Books.
- Gruber, H. E. (1974). Darwin on man: A psychological study of scientific creativity. New York: Dutton.
- Gruber, H. E., & Davies, S. N. (1988). Inching our way up Mount Olympus: The evolving-systems approach to creative thinking. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* (pp. 243–270). Cambridge: Cambridge University Press.
- Gruber, H. E., & Wallace, D. B. (1999). The case study method and evolving systems of approach for understanding unique creative people at work. In R. E. Sternberg (Ed.), *Handbook of creativity* (pp. 93–115). Cambridge: Cambridge University Press.
- Heyes, C., & Hull, D. L. (Ed.) (2001). Selection theory and social construction. Albany: State University of New York Press.
- Holland, J. H. (1975). Adaptation in natural and artificial systems. Ann Arbor, MI: University of Michigan Press.
- Hull, D. L. (1988). Science as a process: An evolutionary account of the social and conceptual development of science. Chicago: Chicago University Press.
- Hull, D. L. (2001). Science and selection: Essays on biological evolution and the philosophy of science. Cambridge: Cambridge University Press.
- Hull, D. L., Glenn, S. S., & Langman, R. E. (2001). A general account of selection: Biology, immunology and behavior. *Behavioral and Brain Sciences*, 24, 511–528.

- James, W. (1880). Great men and their environment. In F. H. Burkhardt, F. Bowers, & I. Skrupskelis (Eds.), *The works of William James, Bd. 6: The will to believe and other essays in popular philosophy* (pp. 163–189). Cambridge, MA: Harvard University Press.
- Kim, K. (2001). Nested hierarchies of vicarious selectors. In C. Heyes & D. L. Hull (Eds.), Selection theory and social construction (pp. 101–118). Albany: State University of New York Press.
- Koza, J. R. (1992). Genetic programming: On the programming of computers by means of natural selection. Cambridge, MA: MIT Press.
- Kronfeldner, M. E. (2007). Is cultural evolution Lamarckian? Biology and Philosophy, 22, 493-512.
- Kronfeldner, M. E. (forthcoming). Creativity naturalized. The Philosophical Quarterly.
- Luria, S. E., & Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, 28, 491–511.
- Mesoudi, A. (2008). Foresight in cultural evolution. Biology and Philosophy, 23, 243–255.
- Mesoudi, A., Whiten, A., & Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, 29, 329–383.
- Mumford, M. D. (1999). Blind variation or selective variation?: Evaluative elements in creative thought. *Psychological Inquiry*, 10, 344–348.
- Mumford, M. D., & Antes, A. L. (2007). Commentaries: Debates about the 'general' picture: Cognition and creative achievement. *Creativity Research Journal*, 19, 367–1374.
- Nickles, T. (2003). Evolutionary models of innovation and the Meno problem. In L. V. Shavinina (Ed.), International handbook on innovation (pp. 54–78). Amsterdam: Elsevier.
- Poincaré, H. (1908). L' invention mathématique. In H. Poincare (Ed.), Science et méthode (pp. 43–63). Paris: Flammurion. (Reprinted in and quotes from *Foundations of science*, 3rd ed., by H. Poincaré, 1982; G. B. Halsted, Trans., pp. 383–394, Washington: University Press of America).
- Popper, K. R. (1972). Objective knowledge: An evolutionary approach. Oxford: Clarendon Press.
- Popper, K. R. (1974a). Intellectual autobiography. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper*, Bd.1. (pp. 3–184). LaSalle: Open Court.
- Popper, K. R. (1974b). Campbell on the evolutionary theory of knowledge. In P. A. Schilpp (Ed.). *The philosophy of Karl Popper* (pp. 1059–65). LaSalle: Open Court. (Reprinted in and quoted from *Evolutionary epistemology, rationality and the sociology of knowledge*, pp. 115–120, by G. Radnitzky & W. W. Bartley, 1987, La Salle: Open Court).
- Popper, K. R. (1975). The rationality of scientific revolutions. In R. Harré (Ed.), *Problems of scientific revolution* (pp. 72–101). Oxford: Clarendon Press. (Reprinted in and quotes from *Scientific revolutions*, pp. 80–106, by J. Hacking, Ed., 1981, Oxford: Oxford UP).
- Popper, K. R. (1984). Evolutionary Epistemology. In J. W. Pollard (Ed.), Evolutionary theory: Paths into the future (pp. 239–255). New York: Wiley.
- Richards, R. J. (1977). The natural selection model of conceptual evolution. *Philosophy of Science*, 44, 494–507.
- Richerson, P. J., & Boyd, R. (2005). Not by genes alone: How culture transformed human evolution. Chicago: University of Chicago Press.
- Roberts, R. M. (1989). Serendipity: Accidental discoveries in science. New York: Wiley.
- Runco, M. A. (2007). *Creativity: Theories and themes, research, development, and practice*. Amsterdam: Elsevier.
- Ruse, M. (1974). Cultural evolution. Theory and Decision, 5, 413-440.
- Schaffer, S. (1994). Making up discovery. In M. A. Boden (Ed.), *Dimensions of creativity* (pp. 13–51). Cambridge, MA: MIT Press.
- Schooler, J. W., & Dougal, S. (1999). Why creativity is not like the proverbial typing monkey. *Psychological Inquiry*, 10, 351–356.
- Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 65–124). Cambridge, MA: MIT Press.
- Simon, H. A. (1981). The Sciences of the Artificial, 2. ed., Cambridge, MA: MIT Press.
- Simonton, D. K. (1988). *Scientific genius: A psychology of science*. Cambridge: Cambridge University Press.
- Simonton, D. K. (1995). Foresight in insight? A Darwinian answer. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 465–494). Cambridge, MA: MIT Press.
- Simonton, D. K. (1999a). Origins of genius: Darwinian perspectives on creativity. Oxford: Oxford University Press.

- Simonton, D. K. (1999b). Creativity as blind variation and selective retention: Is the creative process Darwinian. *Psychological Inquiry*, 10, 309–328.
- Simonton, D. K. (2003). Human creativity: Two Darwinian analyses. In S. M. Reader & K. N. Laland (Eds.). Animal innovation (pp. 309–325). Oxford: Oxford University Press.
- Simonton, D. K. (2005). Darwin as a straw man: Dasgupta's (2004) Evaluation of creativity as a Darwinian process. Creativity Research Journal, 17, 299–308.
- Simonton, D. K. (2007a). Featured article: The creative process in Picasso's Guernica sketches: Monotonic improvements versus nonmonotonic variants. *Creativity Research Journal*, 19, 329–344.
- Simonton, D. K. (2007b). Rejoinder: Picasso's Guernica: Creativity as a Darwinian process: Definitions, clarifications, misconceptions, and applications. *Creativity Research Journal*, 19, 381–394.
- Skagestad, P. (1978). Taking evolution seriously: Critical comments on D.T. Campbell's evolutionary epistemology. *The Monist*, 61, 611–621.
- Sniegowski, P. D., & Lenski, R. E. (1995). Mutation and adaptation: The directed mutation controversy in evolutionary perspective. Annual Review of Ecology and Systematics, 26, 553–578.
- Sober, E. (1992). Models of cultural evolution. In P. Griffiths (Ed.), Trees of life: Essays in philosophy of biology (pp. 17–39). Dordrecht: Kluwer.
- Stein, E., & Lipton, P. (1989). Where guesses come from: Evolutionary epistemology and the anomaly of guided variation. *Biology and Philosophy*, 4, 33–56.
- Sternberg, R. J. (1998). Cognitive mechanisms in human creativity. *Journal of Creative Behavior, 32*, 159–176.
- Sternberg, R. J. (2006). Cognitive psychology (4th ed.). Belmont: Thompson.
- Sternberg, R. J. (Ed.) (1988). The nature of creativity: Contemporary psychological perspectives. Cambridge: Cambridge University Press.
- Thagard, P. (1988). Against evolutionary epistemology. In P. Thagard (Ed.), Computational philosophy of science (pp. 101–111). Cambridge, MA: MIT Press.
- Toulmin, S. (1972). *Human understanding: The collective use and evolution of concepts.* Princeton, NJ: Princeton University Press.
- Ward, T. B., Smith, S. M., & Finke, R. A. (1999). Creative cognition. In R. J. Sternberg (Ed.), Handbook of creativity (pp. 189–212). Cambridge: Cambridge University Press.
- Weisberg, R. W. (1993). *Creativity: Beyond the myth of genius*. New York: Freeman and Company. Wimsatt, W. C. (1999). Genes, memes and cultural heredity. *Biology and Philosophy*, 14, 279–310.