

Creative Outcome as Implausible Utility

J. Y. Tsao¹, C. L. Ting¹, and C. M. Johnson¹

Review of General Psychology
2019, Vol. 23(3) 279–292
© 2019 The Author(s)



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1089268019857929
journals.sagepub.com/home/rgp



Abstract

Two perspectives are used to reframe Simonton's recent three-factor definition of creative outcome. The first perspective is functional: that creative ideas are those that add significantly to knowledge by providing both utility and learning. The second perspective is calculational: that learning can be estimated by the change in probabilistic beliefs about an idea's utility before and after it has played out in its environment. The results of the reframing are proposed conceptual and mathematical definitions of (a) creative outcome as the product of two overarching factors (utility and learning) and (b) learning as a function of two subsidiary factors (blindness reduction and surprise). Learning will be shown to depend much more strongly on surprise than on blindness reduction, so creative outcome may then also be defined as "implausible utility."

Keywords

creativity, utility, learning, blindness, surprise, implausibility

Introduction

Creativity is increasingly regarded as central to human knowledge production (Dietrich, 2015; Hennessey & Amabile, 2010). From human (Guilford, 1967) to machine (Schmidhuber, 2010) learning, from scientific (Feist, 2008) to artistic (Catmull & Wallace, 2014) innovation, from geniuses (Simonton, 1988) to ordinary mortals (Simon, 1977), creativity helps advance the knowledge that enables humanity to adapt to its environments. Better definitions and quantitative assessments of creative outcome might thus accelerate knowledge advances in humans and, perhaps, in cognitive entities more generally. How to define creative outcome with conceptual and mathematical concreteness, however, is still an ongoing challenge (Plucker, Beghetto, & Dow, 2004; Runco & Jaeger, 2012).

Recently, Simonton (2016b) proposed a three-factor definition which led to an intuitive eightfold typology of creative outcome and its "uncreative" variants (Simonton, 2016a). Most importantly, that proposed definition provides a foundation for increasingly concrete definitions of creative outcome. This article builds on and reframes Simonton's definition from two different perspectives.

The first perspective is *functional* (Campbell, 1960; Dietrich, 2015; Simonton, 2011): that, in the exploration and exploitation processes that evolutionary adaptive systems use to produce and employ knowledge about their environments, creative outcome is posited to be successful exploration hence significant advance in that knowledge. Creative ideas are those that add significantly to knowledge by providing both utility *and* learning ("useful learning"):

Ideas that provide utility are valuable and add to knowledge (Popper, 1962/2014), and ideas that also provide learning by changing expectations (probabilistic beliefs) about what is and isn't useful are even more valuable because they force a reshaping of that knowledge.

The second perspective is *calculational*: That the learning part of useful learning can be estimated by the change in expectations (probabilistic beliefs) about an idea's utility before and after it has played out in its environment (Campbell, 1960; Cropley, 2006; Simon, 1977; Simonton, 2011). We imagine the following stylized sequence of events: An initial (prior) assessment is made of the utility of the idea, then the consequences of the idea are played out in its environment, and finally a subsequent (posterior) assessment is made of the utility of the idea. From the change in the two assessments, an estimate can be made of learning: The change in perceived utility before and after the idea has been played out and thus of the degree to which the idea has forced a reshaping of knowledge. As will be described in sections "Functional Perspective: Creativity's Evolutionary Purpose Is Useful Learning" and "Calculational Perspective: Learning as a Change in Probabilistic Beliefs About Utility," learning will be shown to be dominated by surprise, the change in the perceived utility of the idea. Thus, creative

¹Sandia National Laboratories, Albuquerque, NM, USA

Corresponding Author:

J. Y. Tsao, Material, Physical, and Chemical Sciences Center, Sandia National Laboratories, Albuquerque, NM 87185, USA.
Email: jytsoa@sandia.gov

outcome is a combination of high utility and *surprise* at that high utility, or “implausible utility.”

In scientific knowledge, an example of creative outcome and implausible utility might be Alfred Wegener’s (1922/1966) theory of continental drift. Put forth in 1912, the idea was initially disbelieved (given a negative prior assessment). During the subsequent four decades, the consequences of the idea were played out in numerous experiments until the idea was confirmed in the early 1950s (given a positive posterior assessment), ultimately becoming the basis for the modern theory of plate tectonics. In technological knowledge, an example of creative outcome and implausible utility is the laser. First demonstrated in 1960, the laser was widely thought to be “an invention looking for an application” (Constable & Somerville, 2003). Its applications unfolded only gradually, culminating in the 1970s when its revolutionary significance for fiber-optic communications was recognized. In cultural knowledge, an example of creative outcome and implausible utility might be women’s suffrage in the U.S. First proposed seriously in the 1840s but to significant opposition, it was not until 1920 that the 19th Amendment (“the right of citizens of the United States to vote shall not be denied or abridged by the United States or by any State on account of sex”) to the U.S. Constitution was adopted.

Importantly, all examples of creative outcome share the common characteristic of some degree of implausible utility, though they might differ in various ways from these examples. The idea might be larger (Darwin’s theory of evolution) or smaller (a possible car-pool shortcut from the neighborhood to the office) in societal impact, and the idea could be singular or a set of mutually reinforcing ideas—a “paradigm,” in the classic descriptions of such mutually reinforcing ideas in scientific (Kuhn, 1962/2012) or technological (Dosi, 1982) knowledge, or a “schema,” in the broader context of human belief structures (Reisenzein, Horstmann, & Schützwohl, 2019). The gestation period during which the consequences and utility of the idea are “played out” could be more difficult and take longer (involving many knowledge domains and actors) or be simpler and take very little time (involving few knowledge domain and actors). The knowledge domain of the idea might be those occupied by artists, writers, designers, or scientists, with each domain differing in the degree to which the assessments passed on the ideas are grounded in subjective human biases versus objective natural reality and thus in the degree to which creativity is not absolute but can change with time according to the vagaries and historical contingencies of the environment which surrounds it (Csikszentmihalyi, 1996).

The remainder of this article is organized in the following way. Section “Functional Perspective: Creativity’s Evolutionary Purpose Is Useful Learning” discusses the functional perspective of creativity as having the

evolutionary adaptive purpose of successful exploration, hence significant knowledge advance via useful learning. Section “Calculational Perspective: Learning as a Change in Probabilistic Beliefs About Utility” discusses the calculational perspective of estimating learning via the change in expectations (probabilistic beliefs) about an idea’s utility before and after it has played out in its environment. Section “Creative and Uncreative Outcome Typology and Underlying Factors” discusses the resulting typology of creative and uncreative outcomes and connects that typology with more traditional creativity concepts. Finally, note that the term “creativity” is ambiguous and can refer either to creative outcome or to creative process (that probabilistically leads to creative outcome) (Simonton, 2004, p. 15). This article deals almost exclusively with creative outcome, but in section “Implications for Creative Process,” some implications of our proposed definition of creative outcome on creative process are discussed.

Functional Perspective: Creativity’s Evolutionary Purpose Is Useful Learning

Our functional perspective begins with Simonton’s (2013, 2016b) *ansatz* for the creative outcome, c , of an idea:

$$c = u \cdot (1 - p) \cdot (1 - v). \quad (1)$$

The three parameters on the right side of Equation 1 are as follows: u , the eventual utility of the idea; p , the probability that the idea would have been proposed at all (what Simonton called the idea’s “initial response strength”); and v , the prior knowledge (degree of certainty) of the eventual utility of the idea. Each parameter varies continuously between 0 and 1, but when taken to their binary (0 or 1) extremes, define $2^3 = 8$ “pure” types. One of those pure types $(u, p, v) = (1, 0, 0)$ represents an optimally creative outcome, whereas the other seven pure types represent uncreative outcomes—consistent with common intuition that there are many more ways to fail than succeed in being creative, reminiscent of Tolstoy’s “all happy families are alike; each unhappy family is unhappy in its own way” (Simonton, 2016b).

Here, Simonton’s powerful *ansatz* is reframed from a functional perspective: That creativity has an evolutionary adaptive purpose. In the exploitation and exploration processes by which cognitive entities interact with their environments (Hills et al., 2015; March, 1991; Mehlhorn et al., 2015), we posit that creativity’s function is successful exploration that advances knowledge about the world (Campbell, 1960; Dietrich, 2015; Simonton, 2011)—knowledge that in turn enables more successful subsequent exploitation and exploration of the world.

How does an idea advance knowledge? It does so in two distinct ways. The first way is in the intellectual content of the idea itself: *If an idea has high utility, then it adds positively to the existing body of knowledge.* The second way is in how much has been learned from the idea: *If the utility of the idea, after its consequences have been played out, contradicts its utility as initially predicted on the basis of current knowledge, then it forces a reshaping of that knowledge.* Exactly how usefulness and learning alter and reshape knowledge is beyond the scope of this article;¹ here, we simply posit that the greater the utility and learning that have been provided by the idea, the more significant the advance of knowledge. Thus, if creativity’s evolutionary adaptive purpose is successful exploration that advances knowledge, then it is inextricably linked with usefulness and learning, or with what might be called “useful learning.”

How might “useful learning” be characterized mathematically? Here, we propose a simple reframing of Simonton’s powerful Equation 1 ansatz. The first factor in Equation 1, utility, is kept as is because it is common to both framings. The second two factors in Equation 1, however, are not kept as they are. Because both those factors ($1 - p$, related to originality (Simonton, 2013), and $1 - v$, related to Campbell’s (1960) and Simonton’s (2010) blindness) can be thought of as “advance indicators” of learning, l , we make our own ansatz and reframe Simonton’s ansatz for creative outcome to

$$c = u \cdot l. \quad (2)$$

In this reframing, creative outcome is the product of utility and learning, hence a measure of “useful learning.” If either utility or learning is low, creative outcome is also low; for creative outcome to be high, both utility and learning must be high.

To see how the general behavior of $c = u \cdot l$ makes intuitive sense, consider the scenarios in Figure 1 associated with a hunt for a treasure that is in one of six bins. A particular bin is proposed to be searched, and an assessment is made as to how likely that, according to best current knowledge, the treasure will in fact be in that bin. The bin is then searched, the treasure is either found or not, and an updated assessment is made as to whether the treasure was in that bin or not.

Although there are various intermediate scenarios, shown in Figure 1 are just the four extreme scenarios, organized into a 2×2 matrix, associated with whether the result of the search was useful (was the treasure found) and whether the search led to learning (was there a change in the assessment of whether or not the treasure would be in the bin that was searched):

- In the lower left quadrant, the red bin is proposed to be searched, current knowledge *didn’t* expect the

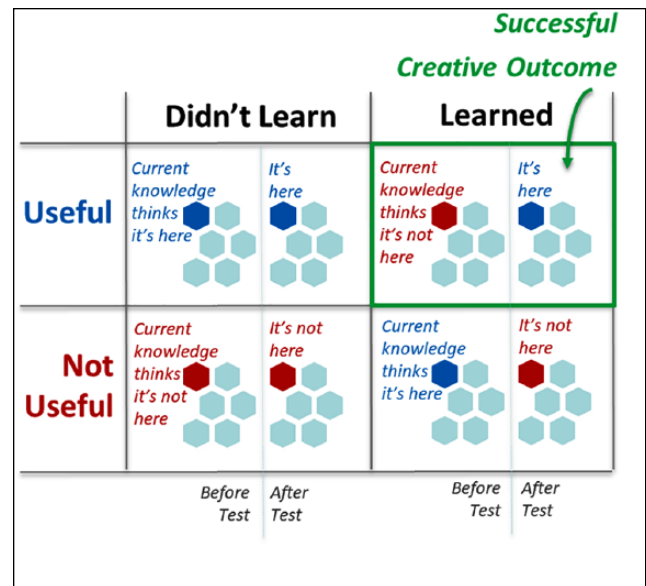


Figure 1. A 2×2 matrix of scenarios associated with a search for a treasure that is in one of six bins.

Note. Bins are proposed to be searched, and an assessment based on current knowledge is made as to whether the treasure will actually be in that bin or not. After the bin is searched, it becomes known whether the treasure is in that bin or not, and whether the assessment of the treasure being in that bin or not was correct or not. Only in the scenario in the upper right quadrant (outlined in green) has “useful learning” occurred: usefulness in the sense that the treasure has been found, and learning in the sense that current knowledge as to whether or not the treasure would be in a particular bin has been overturned.

treasure to be in that bin, and, after the search, indeed the treasure is found *not* to be in that bin. This result is not useful (in that the treasure was not found), and on top of that little has been learned (because current knowledge all along didn’t think that the treasure would be in the red bin).

- In the lower right quadrant, the blue bin is proposed to be searched, current knowledge *expects* the treasure to be in that bin, but, after the search, the treasure is found *not* to be in that bin. This result isn’t useful either, but at least something has been learned (current knowledge has been overturned).
- In the upper left quadrant, the blue bin is proposed to be searched, current knowledge *expects* the treasure to be in that bin, and, after the search, the treasure *is* found to be in that bin. The result is useful (the treasure was found), but little has been learned (because current knowledge all along thought the treasure would be in the blue bin).
- It is the upper right quadrant, outlined in green, where both utility and learning come together. The red bin is proposed to be searched, current knowledge *didn’t* expect (or thinks it implausible) that the treasure would be in that red bin, but, after the search,

the treasure *is* found in that bin. The result is useful (the treasure was found), and on top of that learning has occurred (current knowledge has been overturned). Useful learning and “implausible utility” have both taken place, so creativity, as reframed in Equation 2, is high.

Calculational Perspective: Learning as a Change in Probabilistic Beliefs About Utility

In section “Functional Perspective: Creativity’s Evolutionary Purpose Is Useful Learning,” the *functional* perspective of creativity was discussed. In this section, the *calculational* perspective of creativity is discussed. In this perspective, the learning part of useful learning is estimated by a change in expectations (probabilistic beliefs) about an idea’s utility before and after the idea has played out in its environment.

Assessments of Utility Are Probabilistic

We start with the notion that an idea is associated with a probability *distribution* over utility. As an example, suppose the new idea is for a material that might enable staying warm in a cold environment. Due to incomplete knowledge about the idea before its consequences have been played out in the environment (as well as incomplete knowledge about the environment, which may itself be changing), current knowledge assigns a most probable utility for the idea, but also allows for probabilities that it will have higher or lower utilities. The probability distribution $p(u)$ over utility u is characterized by two properties, as illustrated in Figure 2.

The first property is the mean utility, \bar{u} , which is assumed for simplicity to correspond to the most probable utility. This mean utility is referenced to the set of ideas already known to current knowledge, that is, it is not an absolute value, but rather relative to the state of the art, or more precisely to the current knowledge used to exploit the environment. This case corresponds to Simonton’s “routine expertise” (Simonton, 2016b), to the patent office’s “person having ordinary skill in the art” (U.S. Patent and Trademark Office, 2008), and to the application of current knowledge to a problem routinely and expertly. The reference utility is denoted by the horizontal dashed green line in Figure 2.

The second property is the width (or standard deviation), σ , which is identified with “blindness,” an ingredient essential to *any* evolutionary process, including that for human knowledge. As articulated by Campbell (1960), “a blind-variation-and-selective-retention process is fundamental to all inductive achievements, to all genuine increases in knowledge, to all increases in fit of system to environment.” An idea is blind to the extent that its utility is uncertain, just as an idea is sighted to the extent that its utility is certain (Simonton, 1999).²

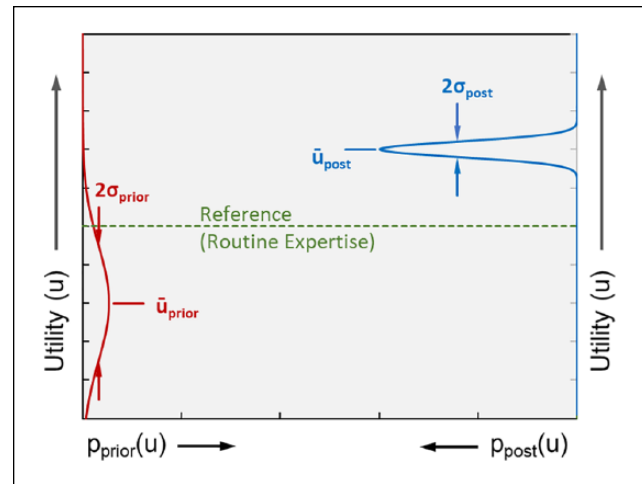


Figure 2. Probability distributions $p(u)$ over the utility (u) of an idea.

Note. On the left, in red, is the distribution on prior assessment according to best current knowledge; on the right, in blue, is the distribution on posterior assessment after the consequences of the idea have been played out. Both distributions are characterized by two parameters: a mean utility, \bar{u} , and a width (or standard deviation), σ . Assumed throughout is that the test is a “good” one so that the width of the probability distribution after test, as depicted, is relatively narrow. The green dashed line represents the utility of routine expertise and calibrates whether an idea is assessed to have higher or lower utility than those ideas already known by “persons having ordinary skill in the art.”

As discussed in section “Introduction,” estimating learning involves making two assessments, prior and posterior. The prior assessment is an assessment according to best initial human knowledge; whereas the posterior assessment is an assessment that incorporates the learning associated with the playing out of the consequences of the idea. Thus, there are *two* probability distributions over utility, associated with these two assessments. Adopting Bayesian terminology, these can be called the prior and posterior distributions, respectively. Each distribution has a mean and width (or standard deviation); as shown in Figure 2, these are denoted ($\bar{u}_{\text{prior}}, \sigma_{\text{prior}}$) and ($\bar{u}_{\text{post}}, \sigma_{\text{post}}$). Note that both these distributions represent guesses about the actual utility, a joint functional property of the idea and the environment. But the first distribution is a guess without benefit of a playing out of the idea in the environment, whereas the second distribution is a guess with benefit of that playing out.

Learning Is Determined by Blindness Reduction and Posterior Surprise

To derive an expression for learning, let us assume for concreteness and ease of mathematical manipulation that the prior and posterior probability distributions over utility are simple normalized Gaussians (Bell Curves):

$$p_{\text{prior}}(u) = \frac{1}{\sqrt{2\pi}\sigma_{\text{prior}}} e^{-\frac{(u-\bar{u}_{\text{prior}})^2}{2\sigma_{\text{prior}}^2}}. \quad (3)$$

$$p_{\text{post}}(u) = \frac{1}{\sqrt{2\pi}\sigma_{\text{post}}} e^{-\frac{(u-\bar{u}_{\text{post}})^2}{2\sigma_{\text{post}}^2}}. \quad (4)$$

The standard information theoretic measure of the learning, l , that takes place as beliefs are revised from a prior probability distribution $p_{\text{prior}}(u)$ to a posterior probability distribution $p_{\text{post}}(u)$ is the Kullback–Leibler (KL) divergence (Cover & Thomas, 2006):

$$l = D_{\text{KL}}(p_{\text{post}}, p_{\text{prior}}) = \int_{-\infty}^{+\infty} p_{\text{post}}(u) \ln \left(\frac{p_{\text{post}}(u)}{p_{\text{prior}}(u)} \right) du. \quad (5)$$

The KL divergence is a measure of the information lost when predicting using the prior rather than the posterior probability distribution and thus is a measure of the information gained when the prior is updated to the posterior probability distribution.

For the Gaussian distributions given above, the KL-divergence learning can be calculated analytically (Baldi & Itti, 2010; Martins, 2013):

$$l = \ln \left(\frac{\sigma_{\text{prior}}}{\sigma_{\text{post}}} \right) - \frac{1}{2} + \frac{1}{2 \left(\sigma_{\text{prior}}^2 / \sigma_{\text{post}}^2 \right)} + \frac{(\bar{u}_{\text{post}} - \bar{u}_{\text{prior}})^2}{2\sigma_{\text{prior}}^2}. \quad (6)$$

If the posterior and prior probability distributions are identical ($\bar{u}_{\text{prior}} = \bar{u}_{\text{post}}$, $\sigma_{\text{prior}} = \sigma_{\text{post}}$), then $l = 0$ and no learning takes place; if they are not identical, then $l > 0$ and some learning takes place. Interestingly, from manipulation of Equation 6, it can be deduced that the magnitude of the learning does not depend on the absolute means or widths of either probability distribution, but instead on two normalized differences between the means and widths. These two normalized differences might be called “blindness reduction” and “posterior surprise.”

Blindness reduction we define to be the change in the widths (or standard deviations) of the prior and posterior probability distributions, normalized to the width of the posterior probability distribution:

$$\Delta b = \frac{\sigma_{\text{prior}} - \sigma_{\text{post}}}{\sigma_{\text{post}}}. \quad (7)$$

When an idea upon prior assessment is as sighted as upon posterior assessment, then the width of the prior probability distribution is the same as that of the posterior probability distribution ($\sigma_{\text{prior}} \gg \sigma_{\text{post}}$), and the blindness reduction is

zero ($\Delta b \gg 0$). When an idea upon prior assessment is much more blind than upon posterior assessment, then the width of the prior probability distribution is much wider than that of the posterior probability distribution ($\sigma_{\text{prior}} \gg \sigma_{\text{post}}$), and the blindness reduction is large ($\Delta b \gg 0$). When an idea upon prior assessment is less blind than upon posterior assessment, the blindness reduction is negative ($\Delta b < 0$), though we do not treat this case here. We are most interested in cases in which an idea is *known* to be creative or not in the fullness of time and hindsight. In other words, we assume “good tests” in that σ_{post} , the posterior blindness, is much narrower than σ_{prior} , the prior blindness. However, we note that the case of posterior probability distributions that are not narrow *is* of interest for other purposes. It gives rise to the possibility of “bad tests”—assessments whose results with respect to high or low mean utility are uncertain enough that best-guess posterior utilities are possibly reversed from actual utilities. In other words, there is the possibility of false positives (high best-guess posterior but low actual utilities) and false negatives (low best-guess posterior but high actual utilities), with ramifications on the long-term accumulation of knowledge.

Posterior surprise we define to be the absolute difference between the means of the prior and posterior probability distributions, normalized by the width of the prior probability distribution:

$$s = \frac{|\bar{u}_{\text{post}} - \bar{u}_{\text{prior}}|}{\sqrt{2}\sigma_{\text{prior}}}. \quad (8)$$

That posterior surprise increases with the difference between the prior and posterior mean utilities is intuitively reasonable: The larger the difference between the prior and posterior mean utility of an idea, the more surprised one is. That posterior surprise is larger the narrower the width of the prior probability distribution is also intuitively reasonable: The less blind one’s guess was, and the more certain one was of the mean utility of the idea, the more surprised one will be if the mean utility of the idea turns out to be different (Faraji, Preuschoff, & Gerstner, 2016). Interestingly but not coincidentally, this definition of posterior surprise is similar to the optimal degree by which knowledge of state variables should be updated in recursive Bayesian estimation (or, in the case of state variables evolving linearly with Gaussian noise, in Kalman filtering) in response to new measurements (Bishop & Welch, 2001; Wikipedia Contributors, 2018b). It is also similar to the surprise discussed by Macedo and Cardoso (2001) in the context of artificial creativity.

Using these definitions for blindness reduction and posterior surprise, Equation 6 for learning can be rewritten as follows:

$$l = \ln(1 + \Delta b) - \frac{1}{2} + \frac{1}{2(1 + \Delta b^2)} + s^2. \quad (9)$$

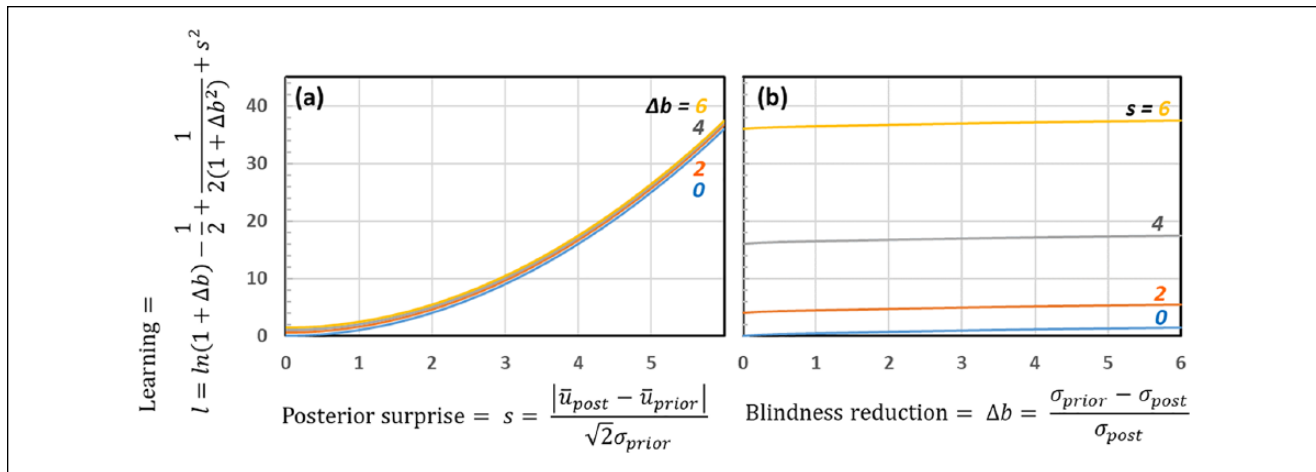


Figure 3. (a) Learning, l , versus posterior surprise, s , for constant values of blindness reduction and (b) learning, l , versus blindness reduction, Δb , for constant values of posterior surprise, s . Learning depends much more sensitively on posterior surprise than on blindness reduction.

The first three terms on the right-hand side of the equation represent learning due to blindness reduction, whereas the fourth (last) term on the right-hand side of the equation represents learning due to posterior surprise. Note that this decomposition of learning into blindness reduction and posterior surprise differs from previous treatments such as Itti and Baldi's (2006) Bayesian surprise. In their treatment, posterior surprise is equated with learning, whereas in this treatment, posterior surprise is one of *two* components of learning. To some extent these are simply mathematical definitions, but, in the context of creativity, we believe it is important to distinguish between blindness reduction and posterior surprise, and so to distinguish their different contributions to learning. It is possible to have blindness reduction that one learns from, but that isn't surprising (we once were never quite knows how the next Moore's Law improvements in integrated circuits will come, but is not surprised when they do come), just as it is possible to have no apparent blindness reduction but be surprised (we once were certain that our local universe is geocentric and then became equally certain that it is heliocentric). Both of these represent learning but of different kinds.

Importantly, as illustrated in Figure 3, learning depends much more sensitively on posterior surprise than on blindness reduction. Learning depends on the square of posterior surprise, but only on the logarithm of blindness reduction (the second and third terms on the right-hand side of Equation 9 become negligible compared with the logarithmic first term at large blindness reduction). In other words, more is learned from changes in the means (\bar{u} s) than from changes in the widths (σ s). In the language of hypothesis testing and the scientific method, more is learned from the quick-and-dirty experiment that *changes* one's view of an idea than from the detailed experiment that simply *confirms*

(though narrowing the uncertainty of) one's view of an idea. To apply this to learning in science, and to borrow Kuhn's language, we might say that surprise is what leads to paradigm shifts and revolutionary science, whereas blindness reduction is what leads to paradigm extensions and normal science (Kuhn, 1962/2012). The magnitude of learning is much larger for the former than for the latter.

Creative and Uncreative Outcome Typology and Underlying Factors

In sections "Functional Perspective: Creativity's Evolutionary Purpose Is Useful Learning" and "Calculational Perspective: Learning as a Change in Probabilistic Beliefs About Utility," a three-factor mathematical definition of creative outcome was proposed. Although the meanings of those factors are somewhat different than in Simonton's definition, we can nonetheless construct a similar typology of creative and uncreative outcomes. In this section, we construct such a typology and then connect it to more traditional creativity concepts.

Underlying Factors

To recapitulate, the three factors that enter into Equations 2 and 9 are as follows. The first factor is blindness reduction (Equation 7), which is proportional to prior blindness (assuming "good tests" and small posterior blindness). Because prior blindness is what intuitively one would mean if one would refer to a just-generated idea as blind, it will sometimes be referred to in shorthand simply as "blindness." The second factor is posterior surprise (Equation 8). Because posterior surprise is what intuitively one would mean if one said one was surprised by how an idea played

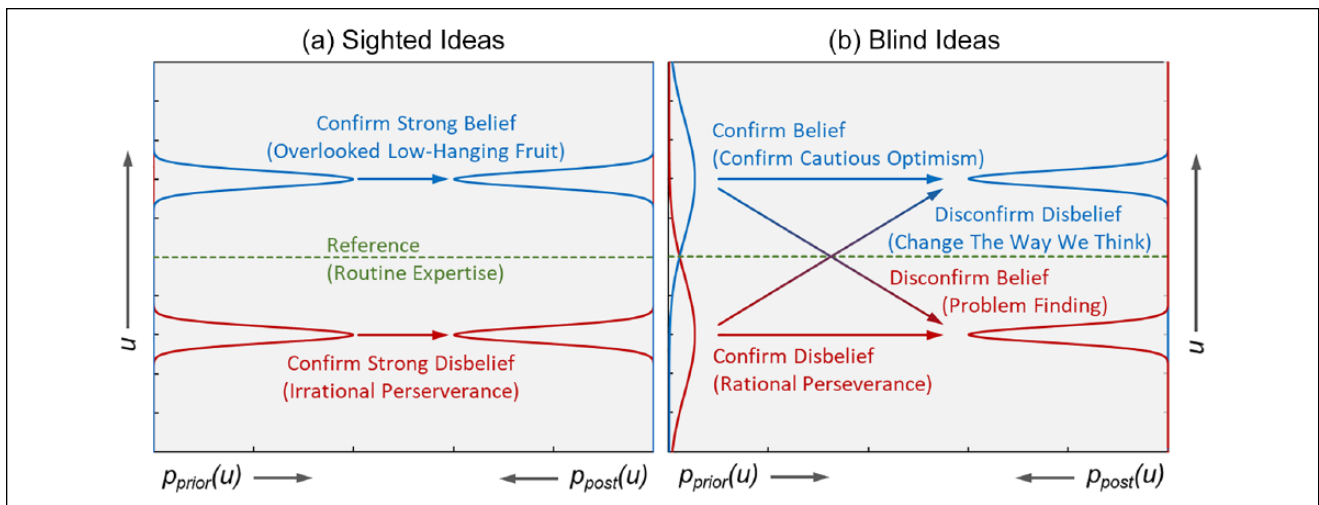


Figure 4. Prior and posterior probability distributions corresponding to two general types of ideas (a) sighted and (b) blind.

Note. In the sighted ideas, on the left, ideas upon generation are more sighted, so more is known about them and their possible utility, and the prior probability distributions over utility are relatively narrow. In the blind ideas, on the right, ideas upon generation are more blind, so less is known about them and their possible utility, and the prior probability distributions over utility are relatively broad. The various types of creativity and uncreativity are determined by the relative positions of the mean utilities before and after the ideas have been played out. There are two types of sighted uncreativity: “confirm strong belief” (“overlooked low-hanging fruit”) and “confirm strong disbelief” (“irrational perseverance”). There are three types of blind uncreativity: “confirm belief” (“confirm cautious optimism”), “confirm disbelief” (“rational perseverance”), and “disconfirm belief” (“problem finding”). There is only one type of creativity: “disconfirm disbelief” (“change the way we think”).

out, it will sometimes be referred to in shorthand simply as “surprise.” The third factor is posterior utility, the same property carried over from Simonton’s typology, and which will sometimes be referred to in shorthand simply as “utility.”

All three of these factors vary continuously, but, for the purpose of a simplified typology, it is assumed in the remainder of this section that they each take on just two extreme values (low and high). In principle, this leads to $2^3 = 8$ types. In practice, however, there are only six types. The reason is that the relationship between prior and posterior utilities is not independent of prior blindness. When prior blindness is low (prior sightedness is high), then the prior and posterior utilities must be similar; when prior blindness is high (prior sightedness is low), then the prior and posterior utilities are free to be dissimilar. Because of this partial correlation, one might say that this is a two-and-a-half-factor typology, rather than a two- or three-factor typology. The resulting six types are summarized in Figure 4 and Table 1. Two are sighted and four are blind, and these are discussed in turn in the following two subsections “Sighted Ideas” and “Blind Ideas.”

Sighted Ideas

The two “sighted” types, illustrated in Figure 4(a), are the “overlooked low-hanging fruit” (high utility) and “irrational perseverance” (low utility) types. These two types are the simplest: sighted ideas with narrow prior probability

distributions over utility. For these types, much is known about the newly generated idea, and the consequences of the idea can be played out intellectually with confidence almost without having to play them out in the world. If the idea is about an alternative route to the grocery store that circumvents an accident on the normal route, with both routes within a neighborhood familiar to current knowledge, current knowledge can be confident about its prior estimate of the utility of the idea.

“*Confirm strong belief*” (“*overlooked low-hanging fruit*”). In the first type, the idea before and after it has been played out has higher mean utility than the reference. There was a strong prior belief that the idea would be useful, and this was confirmed by its playing out. This type might be called “overlooked low-hanging fruit.” This is a “why didn’t I think of that?” idea that was overlooked but, once generated, can be immediately seen to have higher utility than the reference. Note that it had to have been overlooked, because if it hadn’t, it so obviously has higher utility that *it* would be the reference—and indeed, *after* the idea has been incorporated into the knowledge base, it *will* become the new reference, resetting the reference point. An example of this type might be Gauss’ famous trick (Hayes, 2006) for adding the numbers 1 through 100: adding the extreme pair ($101 = 1 + 100$) once and then multiplying the sum by 50. Another example of this type might be Archimedes’ legendary (but possibly apocryphal) “Eureka” bathtub moment when he realized that an object immersed in water displaces a

Table 1. Typology of creative and uncreative outcomes.

Type	Descriptive Rubric	Creativity Typology			
		Intuitive Rubric	Blindness Reduction ("Blindness")	Posterior Utility ("Utility")	Surprise("Surprise")
			$(\sigma_{\text{prior}} - \sigma_{\text{post}})/\sigma_{\text{prior}}$	\bar{u}_{post}	$ \bar{u}_{\text{post}} - \bar{u}_{\text{prior}} / (2^{1/2}\sigma_{\text{prior}})$
Sighted	Confirm Strong Belief	Overlooked Low-Hanging Fruit	L	H	L
	Confirm Strong Disbelief	Irrational Perseverance	L	L	L
Blind	Confirm Belief	Confirm Cautious Optimism	H	H	L
	Confirm Disbelief	Rational Perseverance	H	L	L
	Disconfirm Belief	Problem Finding	H	L	H
	Disconfirm Disbelief	Change the Way We Think	H	H	H

Note. The first five types are uncreative, and the sixth type ("disconfirm disbelief" or "change the way we think") outlined at the bottom is creative. The typology is based on three derived properties of the prior and posterior probability distributions over utility. The three derived properties are: blindness reduction (or "blindness," for short), posterior utility (or "utility," for short), and posterior surprise (or "surprise," for short). For blindness, "H" means relatively blind, "L" means relatively sighted; for utility, "H" means utility higher than the reference, "L" means utility lower than the reference; and for surprise, "L" means relatively low surprise, "H" means relatively high surprise. The descriptive rubrics are suggestive of the mathematical differences between the mean prior and posterior utilities for each type. The rubrics are suggestive of how the types might be viewed intuitively.

volume of water equal to the volume of the object immersed and thus could be used to solve the gold crown "density" problem posed by Hiero of Syracuse. Once one "sees" the idea, and is knowledgeable in the area, it is *obvious* that it will work. But because it is obvious to current knowledge, it is also not likely to overturn it. The degree of utility could very well be profound, but the degree of learning will be less profound.

"Confirm strong disbelief" ("irrational perseveration"). In the second type, the idea before and after it has been played out has lower mean utility than the reference. There was a strong disbelief before it was played out that the idea would be useful, and this was confirmed by its playing out. Following Simonton, this type might be called "irrational perseveration." This is an idea that could be new and overlooked or could be old and have been tried many times before. Once generated, though, it is obvious that it is a "fool's errand," that it will *not* work, or that it will have lower utility than the reference. Thus, it is irrational to persevere in trying the idea. Nonetheless, there are many situations in which people do try ideas that they and current knowledge do not think will work. Examples from daily life might be a bad habit that one knows one should break but cannot, or a superstition that one knows is incorrect but cannot help but follow anyway; or a known bad idea that is tried simply because no better ideas have been found and time is of the essence (a "Hail Mary"). An example from science and technology might be an idea for yet another perpetual motion machine—one that violates energy conservation and therefore could not be possible.

Blind Ideas

The four "blind" types are illustrated in Figure 4(b). For these types, much less is known about the idea as generated, so surprise *is* possible after it has been played out. There are thus four rather than two types, depending on whether utility is high *and* whether one was surprised by that utility.

These four types, interestingly, map approximately and suggestively to four types of reward prediction error (RPE) signals associated with neural representations of an outcome's valence (higher or lower utility than expected) and surprise (deviation from expectations) (Fouragnan, Retzler, & Philiastides, 2018). High utilities are the "confirm belief" and "disconfirm disbelief" types and low utilities are the "disconfirm belief" and "confirm disbelief" types. And, just as in Simonton's typology, in this typology only the last type ("disconfirm disbelief") is identified with creative outcome. It is the one type that combines (prior) blindness, (posterior) utility, *and* (posterior) surprise. All other types might be precursors for future creativity or interesting for other reasons but are in the end uncreative because they are missing one or more of these factors.

"Confirm belief" ("confirm cautious optimism"). In this type, the idea before it has been played out has a high prior utility, and the idea after it has been played out has a high posterior utility. This type might be called "confirm belief," though it could also be called "confirm cautious optimism." It is similar to the "confirm strong belief" type except that there is more uncertainty and thus more opportunity to learn. An example might be the idea that blue-light-emitting diodes

could create economical white light for illumination. The idea was initially thought somewhat plausible but highly uncertain by the technical community, taking about a decade to be proven correct (Haitz & Tsao, 2011). In the limit where uncertainty is extremely large, this type becomes equivalent to Simonton's "fortuitous response," or the "lucky guess." In the limit where uncertainty is extremely small, this type becomes equivalent to the "confirm strong disbelief" or "irrational perseverance" type.

"Confirm disbelief" ("rational perseverance"). In this type, the idea before it has been played out has a low prior utility, and the idea after it has been played out has a low posterior utility. This type might be called "rational perseverance," similar to "irrational perseverance" except that prior blindness is high enough that it might be considered rational to persevere in trying the idea. An example might be cold fusion, which was disbelieved when first proposed, but with just enough uncertainty (as well as enormous potential utility if it were found to be true) to launch world-wide efforts to test and confirm the disbelief (Close, 2014). Another example might be Linus Pauling's 1952 triple helix model for DNA (Pauling & Corey, 1953), a model that was thought unlikely even when it was first published, and that later was indeed proven incorrect by x-ray diffraction data that confirmed Watson and Crick's (1953) alternative 1953 double-helix model. In other words, although the idea is unlikely to have high utility, it might, so thoroughness might dictate persevering with the playing out of the idea.

"Disconfirm belief" ("problem finding"). In this type, the idea before it has been played out has a high prior utility, but the idea after it has been played out has a low posterior utility. This type might be called "disconfirm belief," but, after Simonton, could also be called "problem finding," to highlight the future exploration opportunities it presages. An example might be the Ehrenfest or Rayleigh–Jeans ultraviolet catastrophe: the prediction that an ideal black body at thermal equilibrium should emit more energy at higher frequencies, a prediction that was completely plausible based on the known classical physics of the late 19th century and early 20th century. That prediction did not pass the "test" of conservation of energy, however, and thus could not be correct—indicating a "problem" with the idea that only later was resolved by Planck via quantization, ultimately leading to the development of quantum mechanics (Kuhn, 1987).

"Disconfirm disbelief" ("change the way we think"). In this type, the idea before it has been played out has a low prior utility, but the idea after it has been played out has a high posterior utility. This type might be called "disconfirm disbelief," though it could also be called "change the way we think." An example might be that mentioned in section "Introduction": Alfred Wegener's 1912 theory of continental

drift, initially disbelieved by current knowledge, but then ultimately confirmed and now the basis for the modern theory of plate tectonics (Wegener, 1922/1966). This type represents classic creativity, in which revolutionary ideas that run counter to current knowledge are ultimately proved useful. Flying machines heavier than air, evolution by natural selection, quantum mechanical action at a distance, wave-particle duality, all of these ideas were initially disbelieved but later proved useful and hence "changed the way we think."

Note that the just-previously-discussed type, "disconfirm belief," also can change the way we think, but less directly. To change the way we think, it is not sufficient to find a problem with an idea; the idea with the problem must be followed with an idea that does *not* have a problem, and it is *that* idea that changes the way we think. People typically do not change the way they think even when they know the current way they think is wrong; they only change when there is an alternative way to think that is more satisfying (Kuhn, 1962/2012). Nonetheless, problem finding can certainly be an important stepping stone to creativity, as evidenced by the successes spawned by Thomas Edison's numerous failures (Dyer & Martin, 2010), and as captured in the quote attributed to Isaac Asimov: "The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!' but 'That's funny . . .'"

Connection of Prior Blindness and Posterior Surprise to Traditional Creativity Factors: Novelty/Originality and Nonobviousness

Note that these three "creativity factors"—(prior) blindness, (posterior) utility, and (posterior) surprise—can be taken to be "fundamental" properties derived from the probability distributions. Although utility is a widely used traditional creativity factor, the other two are not. Here, we discuss the two factors, prior blindness and posterior surprise, that are not. In particular, we discuss how these two factors are related to other factors traditionally associated with creativity, including novelty, originality, unlikeliness, unpredictability, and nonobviousness—as well as to one introduced by Simonton (initial response strength).

First, consider prior blindness: An idea has some degree of prior blindness to the extent that, upon generation, current knowledge is uncertain how to predict its utility.

Closely connected to prior blindness is a related factor that might be called "prior surprise." Prior surprise is different from the "posterior surprise" used throughout this article and introduced in section "Calculational Perspective: Learning as a Change in Probabilistic Beliefs About Utility." It is the surprise associated with the idea even being generated at all, even before it is tested, even before it is known whether it has utility or not, and thus even before there could be posterior surprise at that utility. Prior surprise is

the same as Simonton's "initial response strength"—if an idea has a low initial response strength, or low initial probability of being generated at all, then there will be prior surprise if it is generated. And it is somewhat similar to novelty and originality: If an idea is novel or original, it has not been seen before in the context that the cognitive entity is currently addressing, thus is unlikely to be generated in that context (Simonton, 2013), and thus likely to be surprising if it is generated.

Prior blindness, however, is only correlated with, but is not the same as, prior surprise and novelty/originality. If the idea is novel, and hence generates prior surprise, then it is likely not an idea that has been explored in the context of the relevant knowledge domain, and so it is also likely there will be some degree of blindness with respect to the eventual utility of the idea. But an idea could be novel but not blind: Through analogic or other kinds of reasoning, it may be obvious even for an extremely novel idea what its consequences will be (e.g., an idea that violates energy conservation). And an idea might also be not novel but blind: The idea could be a well-known but difficult to prove conjecture that one is blind about (e.g., Fermat's Last Theorem, before it was finally solved in 1995). Thus, prior surprise and, by implication, initial response strength, novelty and originality, are clearly correlated with, and might be used at times as a shorthand heuristic for, prior blindness. But prior blindness is the more foundational factor for defining useful learning and creativity.

Second, consider posterior surprise: An idea has some degree of posterior surprise to the extent that, after test, current knowledge is surprised at the utility it ended up having.

Closely connected to posterior surprise, interestingly, is the first foundational factor, (prior) blindness (or, more precisely, blindness reduction). In other words, these two factors are not independent of each other. At one extreme, when prior blindness is low, posterior surprise is unlikely; this is why, in Figure 4 and Table 1, there are only two "sighted" idea types. And, if posterior surprise is unlikely, learning in turn is also unlikely. At the other extreme, when prior blindness is high, posterior surprise is also unlikely. As discussed in connection with Equation 8, posterior surprise depends inversely on prior blindness—the lower the prior blindness, the more one is surprised when the utility turns out to be different than what was expected. In other words, posterior surprise (hence learning) might be maximized for ideas with *intermediate* prior blindness—blind enough for there to be a reasonable probability of posterior surprise, but not so blind as to reduce the posterior surprise by too much. Thus, there is a probabilistic and nonmonotonic relationship between prior blindness and posterior surprise (Maher, 2010).

Also closely connected to posterior surprise is the "non-obviousness" criterion used by the U.S. Patent Office.

However, nonobviousness is correlated with, but not identical to, posterior surprise. An idea could be nonobvious for two very different reasons. The first reason is that the idea was *so* "blind" that one has almost no idea of the utility of the idea upon generation. The second reason is that the idea wasn't so blind that one didn't have some idea of the utility, but the utility turned out after the idea was played out to be very different from the anticipated utility—in other words, one was "surprised." The first would not give as much surprise and would not lead to as much learning, as the second. Thus, nonobviousness is correlated with posterior surprise and might be used at times as a shorthand heuristic for it. But posterior surprise is the key feature, not nonobviousness.

Also closely connected with posterior surprise is novelty/originality (or, as just discussed, prior surprise). However, as has been discussed recently, novelty/originality (which we take to be synonymous³) are correlated with, but are not identical to, posterior surprise (Barto, Mirolli, & Baldassarre, 2013). In human cognition, novelty/originality might be thought of as associated with events or ideas that are not represented in one's "schema or episodic event memory," whereas posterior surprise might be thought of as associated with events or ideas that "disconfirm expectations or beliefs" (Reisenzein et al., 2019). And, in artificial cognition, there are preliminary indications, from reinforcement learning algorithms which combine extrinsic (*u*) with intrinsic (*I*) rewards, that learning is more strongly correlated with surprise than with novelty/originality (Achiam & Sastry, 2017; Bellemare et al., 2016). For example, that we find that our car door is locked could be not at all novel, but if we find it is locked after we thought we had just clicked the key fob's unlock button it becomes surprising and an occasion for questioning our mental model about key fobs and their function (Barto et al., 2013).

Implications for Creative Process

This article has thus far been devoted to defining creative outcome. But we are cautiously optimistic that these definitions might also enable insights into creative processes aimed at probabilistic production of creative outcomes. Here, we discuss two insights, both associated with criteria for generating ideas that might subsequently be considered creative.

Relative Weighting of Anticipated Utility and Learning

Note first an analogy to evolutionary biology. In biology, organismal variants are generated, and those variants are tested in (and by) their world. Those that survive go on to reproduce, inheriting the original variation but also adding yet new variations. As formalized in Fisher's Fundamental Theorem of Natural Selection, the greater the variance in

properties across organisms within each generation, the faster the rate at which the organismal population evolves (become fitter) from generation to generation. Variation, however, is costly, as most variants are less fit and die before reproducing, so the degree of variance is itself an optimizable and evolvable trait. The more complex and changing the world, the more reason to incur the cost of variance; the simpler and more static the world, the less reason to incur the cost of variance. The optimal rate of evolution or “evolvability” (Pigliucci, 2008) depends on the kind of world the organismal population is embedded in.

The analogy to knowledge and creativity is that idea variants are generated, and these idea variants are tested in the context of the existing body of knowledge as they are played out in the world. The measure of variance here is the degree to which the idea differs from or contradicts the existing body of knowledge, hence the degree to which one might anticipate learning/surprise will take place. Thus, the idea-generation process can be skewed either toward anticipated utility or anticipated learning/surprise. Skewing toward learning/surprise, however, is costly, as most ideas that disagree with current knowledge are wrong and will have low utility. Thus, the optimal degree of creativity or “innovability” (Wagner & Rosen, 2014) depends on the kind of world the cognitive entity is embedded in. The more complex and changing the world, the more reason to incur the cost of learning; the simpler and more static the world, the less reason to incur the cost of learning.

At one extreme, if the risks of not learning are high (because the world is changing fast) and/or if proximate utility signals are sparse and related in a complex and indirect way to ultimate utility (Maher, 2010), then optimal exploration might weight learning more heavily (Burda, Edwards, Storkey, & Klimov, 2018). Indeed, at the extreme of worlds that are complex and changing fast, one might imagine it might be optimal to have almost no utility criterion for filtering out ideas low in the knowledge hierarchy. So long as they satisfy some other heuristic for ultimate utility (e.g., interestingness), ideas might be worth keeping “alive” (Stanley & Lehman, 2015) and exploring further just out of “curiosity” (Berlyne, 1966). At the other extreme, if the risks of not learning are low (because the world is not changing fast) and/or if proximate utility signals are dense and related in a simple and direct way to ultimate utility, then optimal exploration might weight utility more heavily.

Thus, one might imagine generating ideas which maximize an anticipatory and weighted creativity:

$$c_{\text{anticipatory}} = u^{1-\alpha} \cdot l^{\alpha}, \quad (10)$$

where $0 < \alpha < 1$. This is a mathematical form which maintains the product form for creative outcome but weights utility and learning differently, reminiscent of the Cobb–Douglas production function used in macroeconomics (Wikipedia

Contributors, 2018a). If $\alpha = 0$, then u is favored regardless of learning: the conservative, low innovability strategy. If $\alpha = 1$, then l is favored regardless of utility: the aggressive, high innovability strategy.

In other words, α might play the role in knowledge growth that fitness variance plays in Fisher’s fundamental theorem of natural selection—the larger the α , the faster the knowledge grows, albeit at the expense of greater risk, uncertainty, and resource consumption. In a simple, unchanging environment, it might be more optimal for α to approach 0; in a complex, rapidly changing environment, it might be more optimal for α to approach 1. Human cognition might be characterized by a particular weighting (with variations across individual humans) based on the worlds to which it had to adapt during the long-term course of human evolution; an engineered or augmented human cognition might adopt a weighting more appropriate to the current world; and a purely artificial cognition might adopt whatever weighting is appropriate to the world in which humans have embedded it.

Informed Contrariness

Suppose now that $\alpha > 0$, so there is some weighting toward learning. How might ideas be generated with the best possibility of leading to learning? It is one thing to understand creative outcome as an ex-post measure: It is creative outcome measured in *hindsight*, after an idea has played itself out. It is totally another thing to develop *foresight*—the possibility of assessing *before* the idea has played itself out the probability that the idea will represent learning (and creative) outcome *after* it has played itself out (Rietzschel, Nijstad, & Stroebe, 2010). In other words, can a theory of anticipatory surprise, anticipatory learning, and anticipatory creativity (Girotra, Terwiesch, & Ulrich, 2010; Storck, Hochreiter, & Schmidhuber, 1995) be developed? Idea generation and playing out is not cost-free, so improved processes for more accurate, albeit probabilistic, assessment of, for example, research proposals en route to them being executed and played out could potentially improve the productivity of society’s research enterprise.

For the case of researchers proposing ideas to current knowledge gatekeepers (research funders and peer reviewers), this might involve understanding differences between what the researcher knows versus what current knowledge knows and “arbitrating” those differences to enhance research success probabilities.

On the one hand, current knowledge is not always right, and it is precisely when the researcher’s idea is implausible to current knowledge that the potential for creative outcome is greatest. Ideas must be generated with some disregard for current knowledge, as ideas that perfectly reflected current knowledge would be nothing more than deductions inherent in current knowledge itself. Thus, idea generation must contain “what if?” divergences, perpetrated by individuals at

least temporarily flaunting some portion of current knowledge and offering up to it contrarian alternatives for test.

On the other hand, current knowledge is usually correct, so if the researcher is going to go against current knowledge, the researcher had better have good reasons. These reasons might be called “inside knowledge”—knowledge or capabilities that the researcher has that current knowledge doesn’t that make the researcher think the researcher is right and current knowledge is wrong. The researcher is an “informed contrarian,” going against current knowledge but in an informed way so as to reduce the risk of going against current knowledge. A close analogy can be made to venture capitalists choosing which start-ups to back. As Peter Thiel, the well-known Silicon Valley venture capitalist, puts it to entrepreneurs he might invest in: “tell me something that’s true that almost nobody agrees with” (Hof, 2014). Or, to build on Pasteur’s famous saying, one might instead say “contrariness favors the informed mind.”

Note that such an “informed contrariness” idea-generation strategy for successful creative outcome may have interesting implications for the emotional and/or heuristic processes inside an individual researcher’s brain. Divergent (idea generation) and convergent (idea test) thinking may still be the overarching two-step cognitive processes (Beaty, Benedek, Silvia, & Schacter, 2016; Ellamil, Dobson, Beeman, & Christoff, 2012; Emmanuel-Avina et al., 2018; Jung, Mead, Carrasco, & Flores, 2013), but with alternative criteria for the internal heuristic and emotional processes for idea selection. Are there, for example, emotional processes that are correlated with contrariness, where contrariness has both intellectual and social components? Might a contrarian intellectual “aha” moment be accompanied by a sense of social fear in some but social neutrality or even excitement in others? Are there, for example, heuristic processes (or combinations of processes) that are correlated with informed contrariness? Boden discusses three types of heuristic processes for creativity: conceptual-space exploration, concept combination, and conceptual-space transformation (Boden, 2004). Perhaps the first enables “informedness” (i.e., enables the researcher to understand the limits of current knowledge), whereas the second and third enable contrariness (i.e., enables the researcher to generate the new combination concept or conceptual transformation that the researcher immediately knows, if the researcher has internalized current knowledge and its limits, is contrary to that current knowledge).

Also note that creativity is a joint process, involving at least two distinct cognitive entities, for example, researchers proposing ideas to knowledge gatekeepers (research funders and peer reviewers). Thus, gatekeepers also play an essential role in creative processes. On the one hand, gatekeepers determine what ideas are plausible or implausible, so they must have a well-developed understanding of current knowledge. On the other hand, they must be open to

researchers proposing to do the implausible—provided they are convinced the researchers are “informed.” Thus, there is an implied shift in emphasis from gatekeepers assessing plausibility/implausibility to assessing informed contrariness—a strategy more consistent with funding people than with funding projects (Narayanamurti & Tsao, 2018). There is also the implication that creative process is not just to be found in the neuroscience of the individual human brain but also in characteristics of the social and cultural environment that the brain is surrounded by.

Acknowledgments

We acknowledge helpful early comments and suggestions from Dean Simonton, Tim Trucano, Eugene Tsao, Laura Swiler, Venkatesh Narayanamurti, Jessica Turnley, and three anonymous reviewers. We also acknowledge early encouragement and support from Rob Leland, Steve Rottler, Jerry Simmons, and Ben Cook. Remaining errors are of course the responsibility of the authors. This work describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work was partially supported by the Laboratory Directed Research and Development Program at Sandia National Laboratories, a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.

Notes

1. Note, however, that this concept of knowledge reshaping is very similar to the concept of improvement of “absorptive capacity”: both have to do with improving the future ability to predict the utility of new ideas, hence to “absorb” and make use of those new ideas (Cohen & Levinthal, 1989).
2. Note that there can be subtleties associated with various mathematical definitions of blindness and sightedness (Simonton, 2013). Here, we opt for the simplest definition of blindness (the width of the probability distribution), but it will be of future interest to connect the various definitions.
3. Note that there are subtleties associated with various interpretations of novelty, originality, and prior surprise (Simonton, 2013). Here, we opt for the simplest interpretation (that the three are identical), but it will be of future interest to connect the various interpretations.

References

- Achiam, J., & Sastry, S. (2017). *Surprise-based intrinsic motivation for deep reinforcement learning* (arxiv: 1703.01732). Retrieved from <https://arxiv.org/abs/1703.01732>
- Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks*, *23*, 649–666.
- Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or surprise? *Frontiers in Psychology*, *4*, Article 907.
- Beaty, R. E., Benedek, M., Silvia, P. J., & Schacter, D. L. (2016). Creative cognition and brain network dynamics. *Trends in Cognitive Sciences*, *20*, 87–95.
- Bellemare, M., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D., & Munos, R. (2016, June). *Unifying count-based exploration and intrinsic motivation*. Paper presented at the Advances in Neural Information Processing Systems, Barcelona, Spain. Retrieved from <https://papers.nips.cc/paper/6383-unifying-count-based-exploration-and-intrinsic-motivation.pdf>
- Berlyne, D. E. (1966). Curiosity and exploration. *Science*, *153*, 25–33.
- Bishop, G., & Welch, G. (2001). *SIGGRAPH 2001 Course 8: An introduction to the Kalman filter*. Retrieved from <https://courses.cs.washington.edu/courses/cse571/03wi/notes/welch-bishop-tutorial.pdf>
- Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. London, England: Routledge.
- Burda, Y., Edwards, H., Storkey, A., & Klimov, O. (2018). *Exploration by random network distillation* (arxiv: 1810.12894). Retrieved from <https://arxiv.org/abs/1810.12894>
- Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review*, *67*, 380–400.
- Catmull, E., & Wallace, A. (2014). *Creativity, Inc.: Overcoming the unseen forces that stand in the way of true inspiration*. New York, NY: Random House.
- Close, F. (2014). *Too hot to handle: The race for cold fusion*. Princeton, NJ: Princeton University Press.
- Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: The two faces of R & D. *The Economic Journal*, *99*, 569–596.
- Constable, G., & Somerville, B. (2003). *A century of innovation: Twenty engineering achievements that transformed our lives*. Washington, DC: Joseph Henry Press.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). Hoboken, NJ: John Wiley.
- Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal*, *18*, 391–404.
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. New York, NY: HarperCollins.
- Dietrich, A. (2015). *How creativity happens in the brain*. Berlin, Germany: Springer.
- Dosi, G. (1982). Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy*, *11*, 147–162.
- Dyer, F. L., & Martin, T. C. (2010). *Edison: His life and inventions*. Portland, OR: The Floating Press.
- Ellamil, M., Dobson, C., Beeman, M., & Christoff, K. (2012). Evaluative and generative modes of thought during the creative process. *NeuroImage*, *59*, 1783–1794.
- Emmanuel-Avina, G. R., Schunn, C. D., Odumoso, T., Silva, A. R., Bauer, T. L., Crabtree, G. W., . . . Tsao, J. Y. (2018). The art of research: Opportunities for a science-based approach. In E. Subramanian, J. Y. Tsao, & T. Odumoso (Eds.), *Engineering a better future* (pp. 167–186). Berlin, Germany: Springer.
- Faraji, M. J., Preuschhoff, K., & Gerstner, W. (2016). *Balancing new against old information: The role of surprise in learning* (arxiv: 1606.05642). Retrieved from <https://arxiv.org/abs/1606.05642>
- Feist, G. J. (2008). *The psychology of science and the origins of the scientific mind*. New Haven, CT: Yale University Press.
- Fouragnan, E., Retzler, C., & Philiastides, M. G. (2018). Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. *Human Brain Mapping*, *39*, 2887–2906.
- Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea generation and the quality of the best idea. *Management Science*, *56*, 591–605.
- Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, *1*, 3–14.
- Haitz, R., & Tsao, J. Y. (2011). Solid-state lighting: “The case” 10 years after and future prospects. *physica status solidi (a)*, *208*, 17–29.
- Hayes, B. (2006). Gauss’s day of reckoning: A famous story about the boy wonder of mathematics has taken on a life of its own. *American Scientist*, *94*, 200–205.
- Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology*, *61*, 569–598. doi:10.1146/annurev.psych.093008.100416
- Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., & Cognitive Search Research Group. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*, 46–54.
- Hof, R. (2014, February 27). Peter Thiel’s advice to entrepreneurs: Tell me something that’s true but nobody agrees with. *Forbes*. Retrieved from <https://www.forbes.com/sites/roberthof/2014/02/27/peter-thiels-advice-to-entrepreneurs-tell-me-something-thats-true-but-nobody-agrees-with/#235232c15d9a>
- Itti, L., & Baldi, P. (2006). Bayesian surprise attracts human attention. *Advances in Neural Information Processing Systems*, *18*, 547–554.
- Jung, R. E., Mead, B. S., Carrasco, J., & Flores, R. A. (2013). The structure of creative cognition in the human brain. *Frontiers in Human Neuroscience*, *7*, Article 330.
- Kuhn, T. S. (1987). *Black-body theory and the quantum discontinuity, 1894-1912*. Chicago, IL: The University of Chicago Press.
- Kuhn, T. S. (2012). *The structure of scientific revolutions* (4th ed.). Chicago, IL: The University of Chicago Press. (Original work published 1962)
- Macedo, L., & Cardoso, A. (2001, March 21–24). *Creativity and surprise*. Paper presented at the in Proceedings of the AISB’01 Symposium on Creativity in Arts and Science (SSAISB), Brighton, UK.
- Maher, M. L. (2010, August 16–17). *Evaluating creativity in humans, computers, and collectively intelligent systems*. Paper presented at the Proceedings of the 1st DESIRE

- Network Conference on Creativity and Innovation in Design, Aarhus, Denmark.
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2, 71–87.
- Martins, T. G. (2013, July 31). Tag archives: Kullback–Leibler divergence. Retrieved from <https://tgmstat.wordpress.com/tag/kullback-leibler-divergence/>
- Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., . . . Gonzalez, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, 2, 191–215.
- Narayanamurti, V., & Tsao, J. Y. (2018). Nurturing transformative U.S. energy research: Two guiding principles. *MRS Energy & Sustainability*, 5, E10.
- Pauling, L., & Corey, R. B. (1953). A proposed structure for the nucleic acids. *Proceedings of the National Academy of Sciences of the United States of America*, 39, 84–97.
- Pigliucci, M. (2008). Is evolvability evolvable? *Nature Reviews Genetics*, 9, 75–82.
- Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist*, 39, 83–96.
- Popper, K. (2014). *Conjectures and refutations: The growth of scientific knowledge*. London, England: Routledge. (Original work published 1962)
- Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The cognitive-evolutionary model of surprise: A review of the evidence. *Topics in Cognitive Science*, 11, 50–74.
- Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2010). The selection of creative ideas after individual idea generation: Choosing between creativity and impact. *British Journal of Psychology*, 101, 47–68.
- Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24, 92–96.
- Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2, 230–247.
- Simon, H. A. (1977). *Models of discovery: And other topics in the methods of science*. Dordrecht, The Netherlands: D. Reidel.
- Simonton, D. K. (1988). *Scientific genius: A psychology of science*. Cambridge, UK: Cambridge University Press.
- Simonton, D. K. (1999). Creativity as blind variation and selective retention: Is the creative process Darwinian? *Psychological Inquiry*, 10, 309–328.
- Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*. Cambridge, UK: Cambridge University Press.
- Simonton, D. K. (2010). Creative thought as blind-variation and selective-retention: Combinatorial models of exceptional creativity. *Physics of Life Reviews*, 7, 156–179. doi:10.1016/j.phrev.2010.02.002
- Simonton, D. K. (2011). Creativity and discovery as blind variation: Campbell's (1960) BVSR model after the half-century mark. *Review of General Psychology*, 15, 158–174.
- Simonton, D. K. (2013). Creative thought as blind variation and selective retention: Why creativity is inversely related to sightedness. *Journal of Theoretical and Philosophical Psychology*, 33, 253–266.
- Simonton, D. K. (2016a). Creativity, automaticity, irrationality, fortuity, fantasy, and other contingencies: An eightfold response typology. *Review of General Psychology*, 20, 194–204.
- Simonton, D. K. (2016b). Defining creativity: Don't we also need to define what is not creative? *The Journal of Creative Behavior*, 52, 80–90.
- Stanley, K. O., & Lehman, J. (2015). *Why greatness cannot be planned: The myth of the objective*. Basel, Switzerland: Springer Science + Business Media.
- Storck, J., Hochreiter, S., & Schmidhuber, J. (1995, October). *Reinforcement driven information acquisition in non-deterministic environments*. Paper presented at the Proceedings of the International Conference on Artificial Neural Networks, Paris, France.
- U.S. Patent and Trademark Office. (2008). *2141 examination guidelines for determining obviousness under 35 U.S.C. 103*. Retrieved from <https://www.uspto.gov/web/offices/pac/mpep/s2141.html>
- Wagner, A., & Rosen, W. (2014). Spaces of the possible: Universal Darwinism and the wall between technological and biological innovation. *Journal of The Royal Society Interface*, 11(97), 20131190.
- Watson, J. D., & Crick, F. H. (1953). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 171, 737–738.
- Wegener, A. (1966). *The origin of continents and oceans* (J. Biram, Trans., 4th ed.). New York, NY: Dover. (Original work published 1922)
- Wikipedia Contributors. (2018a, October 18). *Cobb-Douglas production function*. Wikipedia: The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Cobb-Douglas_production_function
- Wikipedia Contributors. (2018b, August 25). *Recursive Bayesian estimation*. Wikipedia: The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Recursive_Bayesian_estimation