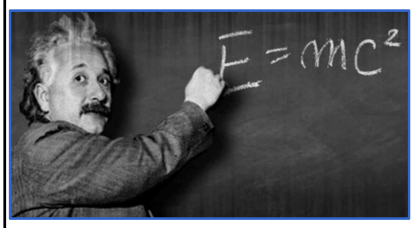
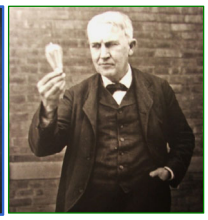
Creativity in Scientific and Engineering Research

Jeff Tsao · Sandia National Laboratories





<u>Why we should care</u>:
Creativity is associated with the greatest advances in human knowledge, how could we not care?
<u>Why we shouldn't care</u>:
Creativity is too difficult a topic, better just "to do" and not overanalyze the "doing"…

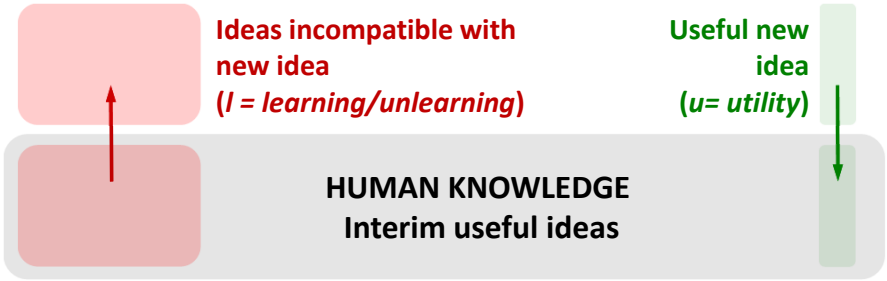
Tsao, J.Y., Ting, C.L. and Johnson, C.M., 2019. Creative outcome as implausible utility. Review of General Psychology, 23(3), pp.279-292.

1. Creative Outcome: *Creative Destruction, Useful Learning, Implausible Utility*
2. Creative Process: *Research Arbitrage, Informed Contrariness*

Sandia National Laboratories is a multimission 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.

- Thanks, Brian, for the invitation to speak on this topic of creativity in scientific and engineering research. Before we dive in, though, let's first ask: why should (or shouldn't) we care about this topic?
  - Why we should care: because creativity is associated with the greatest advances in human knowledge. If we're interested in advancing human knowledge, how could we *not* care about creativity?
  - Why we shouldn't care: because creativity is obviously a difficult nut to crack. It's been studied for a long time, by many very smart people, seemingly without much quantitative progress. Like many things in life, maybe it's better just "to do" and not overanalyze the "doing."
  - Both arguments are valid -- but, for better or for worse, we've been swayed more by the first than the second. I personally feel the stakes are extremely high. If we can accelerate advance in human knowledge, how would that not be incredible for the nation and for the world?
- So starting a few years ago we embarked on a deep dive into this topic. It took us a while, but one of the things we eventually realized is that there are two halves to creativity that one must treat separately: there's creative outcome, and then there is creative process intended to probabilistically lead to creative outcome. Most of what I'll talk about today will be the first half – creative outcome – and will be based on this article that came out last year in Review of General Psychology. But, now that we've made some progress on that first half, we've started to think about the second half – creative process – so I'll also talk a bit about that.
- As a quick preview, some of the phrases you will hear as I talk about creative outcome will be creative destruction, useful learning, and implausible utility; and some of the phrases you will hear as I talk about creative process will be research arbitrage and informed contrariness. If, by the end of the talk, I've conveyed what I mean by those phrases, then I'll have conveyed what I think creativity is.

Our 1st Ansatz: Creative Outcome = Utility *and* Learning

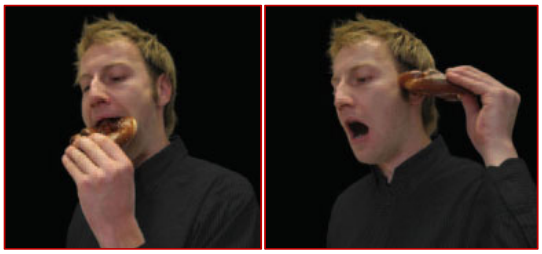


What gets us into trouble isn't what we don't know.
It's what we know for sure that just ain't so.
– (not) Mark Twain

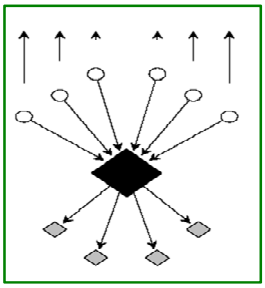
- OK, let's dive in, starting with creative outcome. Imagine that this grey box represents our current stock of human knowledge. It's a collection of interim useful ideas – ideas that we accept as true.
- Now, imagine someone has a new idea that has proven to be useful. Let's say it is for a fabric that is impermeable to water but permeable to air, so it makes great raincoats. We add that idea to our collection of interim useful ideas, and in that way have advanced human knowledge by some amount we might call *u*, the utility associated with the new fabric.
- Now imagine two possibilities.
  - Possibility one: the new idea is totally compatible with the ideas already in our collection of interim useful ideas. If you were to ask a fabric chemist, he or she would say, oh, of course that's how that fabric should perform. There's no need to remove any ideas from our collection of interim useful ideas, no need to do any restructuring of our collection of interim useful ideas. The extent of the advance in knowledge is simply the utility of the new idea.
  - Possibility two, the one indicated here: the new idea *isn't* compatible with some of the ideas in our collection of interim useful ideas. If you were to ask a fabric chemist, he or she would say, oh, that fabric *shouldn't* be impermeable to water, it should be *permeable*. Of course, we've already shown that the fabric makes great raincoats, so that means there's something wrong with some of our ideas about fabric chemistry. We need to remove those ideas from our collection of interim useful ideas, or at least restructure them in some way.
- The advance in knowledge is now more complicated. As a consequence of the new idea, we've had to restructure existing knowledge, by some amount we might call *l*, the learning or unlearning associated with the fact that the new fabric makes great raincoats. And our ansatz is that this learning or unlearning is a very big deal. Nearly every really significant advance in human knowledge has involved a restructuring of existing knowledge in some way, has caused a change in the way we think over and above just the utility of the idea itself. It's a nod to how wrong or at least incomplete are the ideas in our collection of interim useful ideas. We believe all sorts of things that in the end aren't true. It's like the saying that is apparently mis-attributed to Mark Twain: what gets us into trouble isn't what we don't know; it's what we know for sure that just ain't so.

The idea of "creative destruction" isn't new

Gestaltists	1900's	Psychology	Gestalt Shift
Joseph Schumpeter	1940's	Economics	Creative Destruction
Thomas Kuhn	1960's	History of Science	Paradigm Shift
Clay Christensen	1990's	Business	Disruptive Innovation
Laurent Itti & Pierre Baldi	2000's	Psychology	Surprise and Human Learning
Russell Funk	2010's	Bibliometrics	Citation Annihilation



Moritz Koster, Miriam Langeloh and Stefanie Hoehl, "Visually entrained theta oscillations increase for unexpected events in the infant brain," Psychological Science 1-8 (2019).



Russel Funk, Jason Owen-Smith, "A dynamic network measure of technological change," Management Science 63, 791-817 (2017).

- Now, the idea that really significant advances in human knowledge have two halves to them – the creation of something new and the destruction (or at least restructuring) of something old – isn't itself new. Here I list different ways in which that idea has cropped up in various eras and fields.
- There's the Gestaltists who talked about Gestalt shifts. For example, many of you are familiar with this old woman young woman picture. You start by seeing one – either the old woman or the young woman. Then, you try to see the other, and after you've made the Gestalt shift to seeing the other, you can no longer see the first. To see the old woman you must destroy the young woman, or vice versa.
- There's the economist Joseph Schumpeter, who in the 1940's and 1950's coined the phrase creative destruction to describe what he viewed as the great overarching fact of free-market economies: the creation of new ways of doing things accompanied by the destruction of old ways of doing things.
- There's the historian of science Thomas Kuhn, who in the 1950's and 1960's coined the phrase paradigm shift. Once you see the solar system as heliocentric, your view of the solar system as geocentric gets obliterated.
- There's the business economist Clay Christensen, who coined the phrase disruptive innovation to call attention not just to innovation but to the disruption that innovation causes to existing ways of doing things.
- In psychology, we now know that even infants are attracted to events which surprise them, and which cause them to "rethink" what they thought they knew. If you put a bagel in your ear instead of in your mouth, even nine-month-old infants will stare.
- In bibliometrics, the phenomenon of citation annihilation is getting recent press. Papers that are considered significant advances in knowledge screen papers that they cite from papers that they are cited by – in a sense destroying those previous ideas. After Newton, there was no need to cite Aristotle.

Our 2nd Ansatz: Utility and Learning = Useful Learning

Simonton, D. K. (2016). Creativity, Automaticity, Irrationality, Fortuity, Fantasy, and Other Contingencies: An Eightfold Response Typology. *Review of General Psychology, 20*(2), 194-204.

Dean Simonton's Ansatz

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

Novelty Blindness

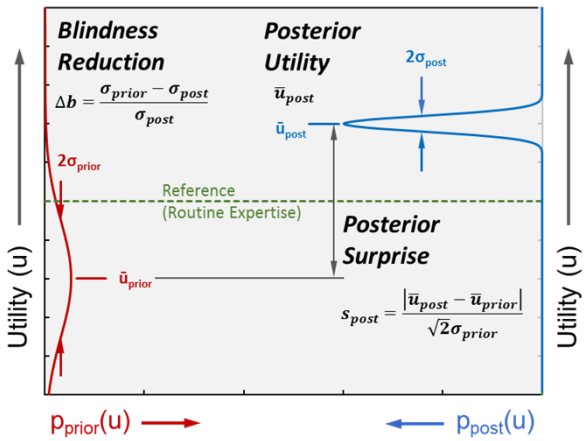
Useful Learning

"Leading indicators" of learning

- OK, let's suppose that creative outcome does involve two halves: the creation of something new and useful, call that utility; and the destruction of what you thought you knew before, call that learning. How might we combine those two halves into one mathematical expression for creativity?
- Fortuitously, Dean Simonton, perhaps this century's pre-eminent creativity theorist, proposed in this paper from a few years ago a new mathematical formulation for creative outcome. In his formulation creative outcome is the product of three factors. The first factor is the utility of the idea after the idea has been played out. The second factor is the novelty of the idea when it was first proposed. The third factor is our lack of knowledge, or our blindness regarding the final utility, of the idea, when it was first proposed. He made the expression a product so that all three factors would need to be high in order for creative outcome to be high. In fact, this formulation basically reproduces what the US Patent Office views as worthy of a patent: it must be useful, it most be novel, and it must not be obvious to those expert in the field.
- I want to emphasize, though, that there is no deeper basis for this equation other than the desire for all three factors to need to be high in order for creative outcome to be high. But we decided to see where we would get if we built on this equation, with one major modification. These two factors "look" a bit like learning, or at least are leading indicators of some kind for learning. If the idea is novel and you are relatively blind at the outset as to whether the idea will be useful or not, then there is a greater chance that you will be surprised at the outcome and will have learned something. So our second ansatz is to replace those two factors with a single factor, learning, giving this simpler expression: creative outcome is the product of utility and learning, something that we might call useful learning.
- Now, of these two halves, utility and learning, the utility half seems intuitive, so we won't discuss that any further. But the learning half isn't so obvious, so let's dig a little more deeply into that half.

Our 3rd Ansatz: Learning = Degree of Knowledge Restructuring

Kullback–Leibler (KL) divergence over prior and posterior assessments of utility



$$l = D_{KL}(p_{post}, p_{prior})$$

$$= \int_{-\infty}^{+\infty} p_{post}(u)\, ln\left(\frac{p_{post}(u)}{p_{prior}(u)}\right) du$$

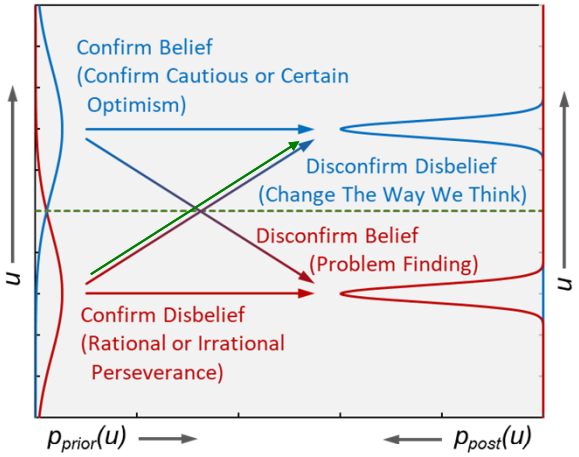
$$= ln(1 + \Delta b) - \frac{1}{2} + \frac{1}{2(1+\Delta b^2)} + s_{post}^2$$

$$\sim s_{post}^2$$

= Implausibility · Implausibility

- OK, what about learning? I've said that learning might be viewed as a changing or restructuring of existing knowledge when that existing knowledge is contradicted by a useful new idea. What might be a measure of the degree of restructuring required? Here we make our third ansatz: that the degree of restructuring required is the degree to which existing knowledge predicted wrongly how the utility of the idea ended up playing out. If existing knowledge predicts perfectly how the utility of the new idea ended up playing out, then no restructuring is necessary. But if existing knowledge predicts badly how the utility of the new idea ended up playing out, a lot of restructuring will be necessary. That's illustrated here.
  - On the left, in red, we plot a possible probability distribution over utility that existing knowledge guessed after a new idea was proposed, but prior to the idea being played out. The left axis is utility, and the bottom axis is the guessed probability that the idea will have that utility. I've drawn the probability distribution to be somewhat broad, and also at the lower end of the utility scale.
  - On the right, in blue, we plot a possible probability distribution over utility posterior to the idea being played out. After the idea has played out, you know its utility, so the probability distribution is narrower, and in this case I've drawn it so that it is at the higher end of the utility scale.
- To measure how much has been learned in going between these two distributions, we use the so-called Kullback-Leibler (KL) divergence over the prior and posterior assessments of utility, where the KL divergence is the standard information-theoretic measure of the information gained when the prior is updated to the posterior probability distribution. I won't go through the details, but if you follow the math through with these two Gaussians, you end up with this expression.
  - Δb is what we call blindness reduction. It's the relative change in the widths of the distributions – the width of the prior distribution minus the width of the posterior distribution, divided by the width of the posterior distribution. In this case the posterior distribution is much narrower than the prior distribution, so the blindness reduction in this case is significant.
  - s_{post} is what we call posterior surprise. It's the change in the best-guess mean utility, normalized to the width of the prior distribution: u_{post} minus u_{prior}, all divided by the width of the prior. In other words, the bigger the change in your guessed utility before and after the idea has played out, the more you are surprised; and the more certain you were of the utility before the idea has played out, the narrower the prior distribution, also the more you are surprised.
- So learning has both of these components: blindness reduction and posterior surprise. But, and this is important, the dependence on blindness reduction is in this log term, whereas the dependence on posterior surprise is in this squared term, so learning is way more sensitive to posterior surprise than it is to blindness reduction. So we can drop the blindness reduction piece, and say that learning is approximately the square of posterior surprise. Because posterior surprise is what we might also call implausibility – it is when it was implausible that the idea be useful but it ends up being useful that you are surprised – we can think of this as the square of implausibility.

Creative and uncreative outcome: fourfold typology



Creative Outcome = ***Useful Learning***

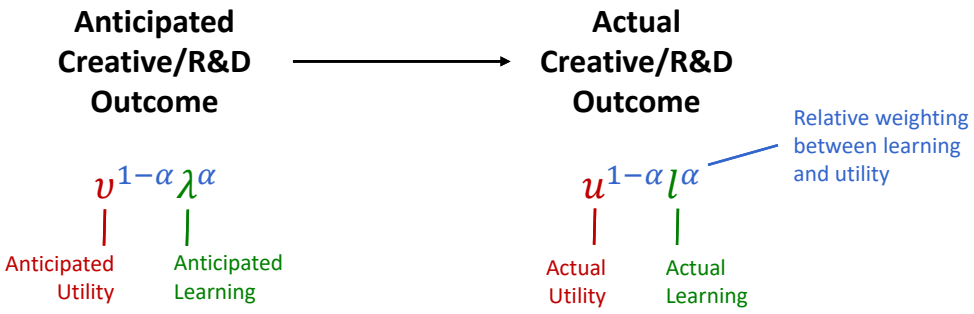
$= u \cdot l$

$\sim u \cdot s^2$

= ***Implausible Utility***

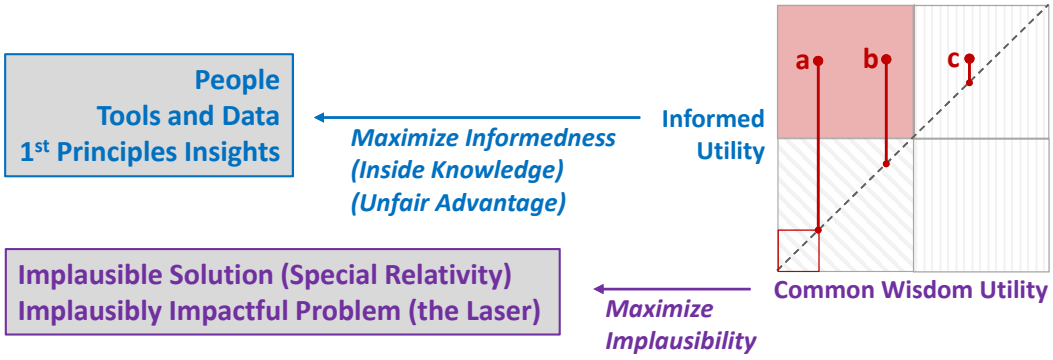
- OK, that's learning. Now let's put learning back together with utility and recapitulate.
  - We started out by saying that creative outcome has two halves to it. There's the idea itself, which has utility. Then there's the restructuring of existing knowledge, which is learning. Really significant knowledge advance involves both utility and learning, so we have useful learning. A simple expression for useful learning is u times l, which has the right limiting behavior – both utility and learning must be high in order for creative outcome to be high.
  - A proxy for how much learning takes place, how much knowledge restructuring is required, is the degree to which existing knowledge mis-predicts the actual utility of the idea. Mis-prediction of utility might be because you were just blind, or ignorant, but it also might be because nature surprised you. The degree of learning due to the decrease in blindness or surprise turns out to be dominated by surprise, as the square of surprise, so we end up with utility times surprise squared.
  - Finally, because surprise in the case where you end up with high utility is the equivalent of implausibility, we end up with something we might call implausible utility.
- And now we have a fourfold typology of creative and uncreative outcome.
  - One type is down here. Current wisdom didn't believe the idea would be useful, and it turns out the idea wasn't useful. You've confirmed your disbelief. This might be a rational thing to do if your initial blindness were high, in which case it might be called rational perseverance. But it might be an irrational thing to do if your initial blindness were low, and it might then be called irrational perseverance.
  - Another type is up here. Current wisdom believed the idea would be useful, and it turns out the idea is useful. You've confirmed your belief. If your initial blindness were high, we might call this confirm cautious optimism, but if your initial blindness were low, we might call this confirm certain optimism.
  - Another type is this one going diagonally downwards. Current wisdom believed the idea would be useful, but it turns out the idea isn't useful. You've disconfirmed belief, you've found a problem in your beliefs and need to learn so that you can update your beliefs. This isn't creative outcome per se, because the idea itself didn't turn out to be useful, but it can be an important precursor to creative outcome, as it can steer one towards ideas that are useful.
  - The final type is the one going diagonally upwards, in green. This is the only one of the four types that represents creative outcome. Current wisdom didn't believe the idea would be useful, and it turns out the idea is useful. You've disconfirmed your disbelief, and in doing so have forced a change in the way we think, in a positive way. The idea was implausible *and* useful.

Creative Process = *Anticipatory* Creative Outcome



- OK, we've spent a fair amount of time on creative outcome. Let's switch gears now and talk about creative process. For this, we must work backwards, asking: suppose we have a metric for *actual* creative outcome; what would be a metric for *anticipated* creative outcome?
- What metric should we use for actual creative outcome? Actual utility and actual learning, of course, but here we introduce an important twist. We aren't just interested in creative outcome, we're interested more generally in R&D outcome. But, as we all know, R is not the same as D. With R there is much more emphasis on exploring the unknown; with D there is much more emphasis on exploiting the known. To take this into account, we introduce a weighting parameter α: we give actual utility an exponent 1-α, and actual learning an exponent α. If the R&D is more D like, one would set α closer to 0 so that utility is more valued than learning; if the R&D is more R like, one would set α closer to 1, so that learning is more valued than utility. This way, an investor in R&D can choose the degree to which he or she wishes to support R versus D, while using the same underlying metrics for utility and learning.
- For *anticipated* creative or R&D outcome we use exactly the same equation with the same α weighting factor. Except now anticipated R&D outcome is anticipated utility, υ, and anticipated learning, λ. Just like before for actual utility, anticipated utility is somewhat intuitive so we're not going to discuss it much. But anticipated learning is not so intuitive, so let's dive a bit more deeply into that.

Anticipated Learning = Research Arbitrage = Informed Contrariness

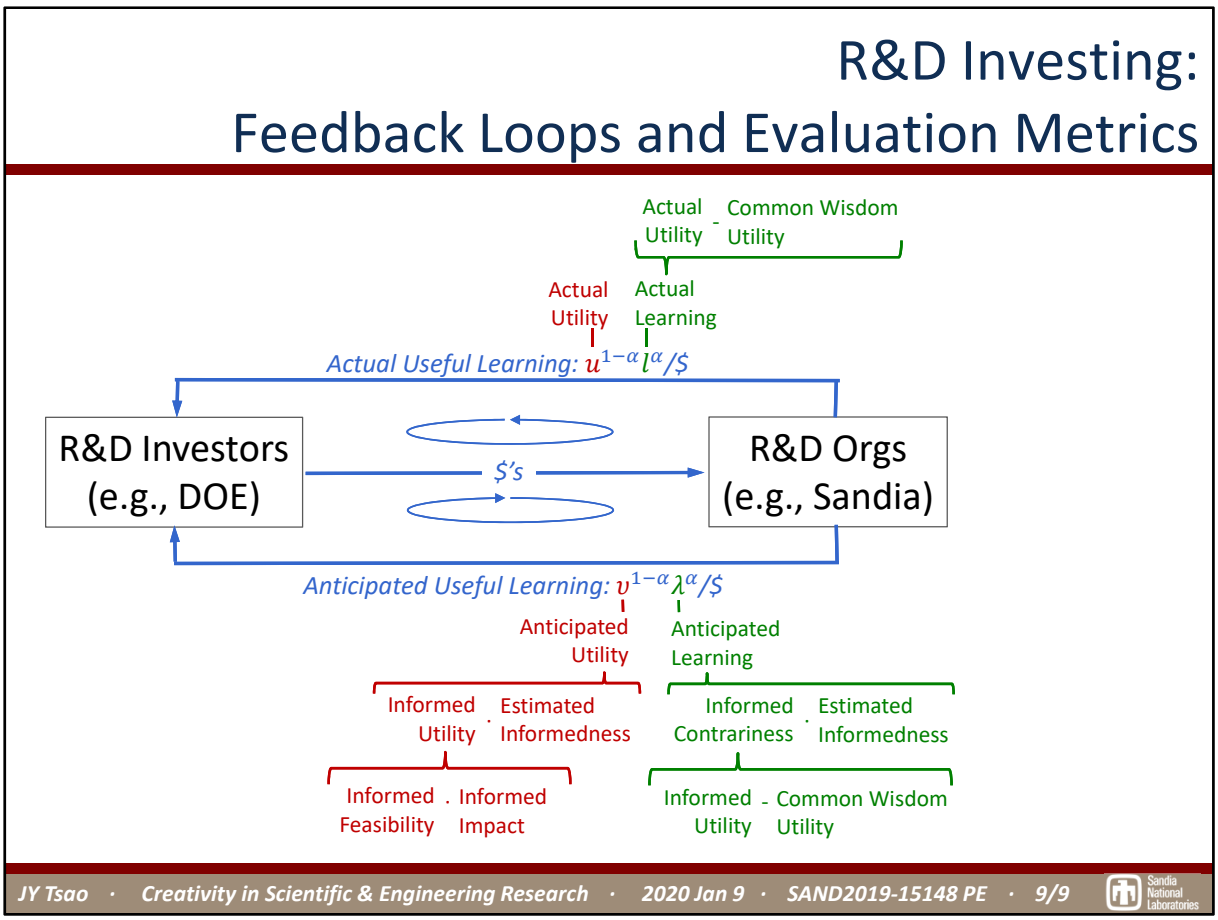


Tell me something that's true that almost nobody agrees with
-- Peter Thiel

Anticipated $ Gain = **Financial Arbitrage** = (*Informed price* - *Common wisdom price*) · **Estimated Informedness**

Anticipated Learning = **Research Arbitrage** = (*Informed utility* - *Common wisdom utility*) · **Estimated Informedness**

- OK, let's dive more deeply into anticipated learning. To do this, I've drawn in the upper right a diagram with two axes.
  - The bottom axis is what common wisdom thinks the utility of the R&D will be before the R&D was done. Let's say that guess is low. Then the R&D was implausible, and somehow went against common wisdom. Why might this be? It might be because there doesn't seem to be any viable approach to the solution of a problem, as when Einstein was working on his theory of special relativity. It could be because the solution seems unlikely to solve an impactful problem, as when Townes and Schawlow were working on the laser. Or it could be both.
  - The left axis is what the informed R&D investor thinks the utility of the R&D will be, again before the R&D was done. If it's low, then the R&D investor agrees with common wisdom. But if it's high, then the R&D investor disagrees with common wisdom. Why might that be? It might be because the R&D investor has some inside knowledge that gives him or her an unfair advantage over common wisdom. Maybe he or she knows that the R&D performer is truly an Einstein or Edison. Maybe he or she knows that the R&D performer has a new tool, a better telescope, that common wisdom doesn't know about yet. Or maybe he or she knows that the R&D performer has thought through the problem more deeply, down to first principles, than common wisdom has. In other words, the R&D investor has an informedness that gives him or her the confidence to predict an overturning of common wisdom.
- And, the greater the discrepancy between the informed utility and the common wisdom utility the greater the anticipated learning. The proposed R&D represented by a, b and c all have the same informed utility as guessed by the R&D investor. But idea a has the lowest common wisdom utility, and therefore has the greatest opportunity for anticipated learning.
  - This if very similar to what happens in financial markets, where we would call it financial arbitrage. If an inside trader's informed knowledge of the price of an asset is higher than common wisdom's knowledge of the price of that asset, then he or she can use that to his or her advantage by investing in the asset. Of course, the anticipated $ gain is moderated by an estimated degree of informedness, or the degree to which the informedness is trustworthy, which is this "estimated informedness" moderating factor here.
  - In research markets, this would be like doing research arbitrage as an informed contrarian – the informed contrarian's knowledge of the value of the proposed R&D is better than common wisdom's knowledge of the value of the proposed R&D, and he or she can use that to his or her advantage by investing in the R&D when no one else will. And, again, the anticipated learning is moderated by an estimated degree of informedness, or the degree to which the informedness is trustworthy, which again is this "estimated informedness" moderating factor here.
- The bottom line is: anticipated learning is maximized when there is the greatest discrepancy between informed and common wisdom utility, moderated by estimated informedness. It's a bit like what Peter Thiel, the high tech venture capitalist says: Tell me something that's true that almost nobody agrees with.



- OK, now let's put it all together. We have actual useful learning, and we have anticipated useful learning – which we can now use as evaluation metrics in various R&D investment strategies.
- To see this, imagine two entities. The entity on the left corresponds to R&D investors. The entity on the right corresponds to R&D execution organizations. There are two possible feedback loops connecting these two entities.
- In the feedback loop at the top, R&D investors supply funding to the R&D organizations, and in turn R&D organizations produce research, in the form of useful learning, for the R&D investors. The more useful learning the more funding, the less useful learning the less funding. The evaluation metrics are relatively simple: actual utility and actual learning, with actual learning being actual utility minus common wisdom utility. In principle this is the best feedback loop, because it rewards *actual* R&D outcome.
- In the feedback loop at the bottom, R&D organizations supply proposals to R&D investors, in the form of anticipated useful learning and, in turn, R&D investors supply funding to the R&D organizations. The more useful learning is anticipated the more funding, the less useful learning is anticipated the less funding. The evaluation metrics here are less simple: anticipated utility and anticipated learning. Anticipated utility is informed utility times estimated informedness; and informed utility, in turn, is something we might call informed feasibility times informed impact. Anticipated learning is informed contrariness times estimated informedness; and informed contrariness, in turn, is informed utility minus common wisdom utility.
- In practice, R&D investors often make two mistakes. First, they rarely use the top feedback loop, even though it is much more direct. Instead, they use the bottom feedback loop, even though it is much more indirect. Second, when they use the bottom feedback loop, they rarely use anticipated utility and anticipated learning. Instead, they mostly just use anticipated utility – that is, they take the $\alpha = 0$ limit. In other words, there is a lot of room for improvement in how R&D investors fund R&D organizations for maximum productivity. Understanding how we might implement those improvements is definitely of great interest, but that's work in progress.
- With that, thank you very much for your attention; happy to answer questions.