SAND2018-7340C

Welcome to Our Symposium!

Creativity Science: Principles and Application to the Funding of Research

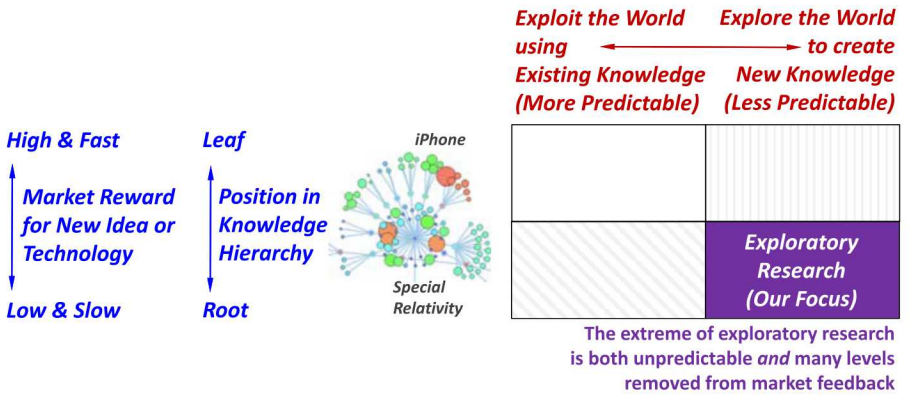
When	How Long	Who	What
1:00p	5m	Jeff Tsao	Overview of Symposium
1:05p	25m	Ken Stanley	How Creativity in Artificial Systems Informs Fostering Human Creativity
1:30p	5m	Curtis Johnson	Audience Q&A
1:35p	10m	Jeff Tsao	Creativity as Research's Meta-Objective
1:45p	5m	Curtis Johnson	Audience Q&A
1:50p	30m	Curtis Johnson	Discussion
2:20p			Adjourn

Jeff Tsao · APS Creativity Symposium · 2018 May 25

- Good afternoon, and welcome to our symposium on "creativity science: principles and application to the funding of research." I'm Jeff Tsao, the organizer of the symposium.
- The intellectual content of the symposium is exactly as it is in the program, but we did have a last-minute change our program. Originally, Julia Klebanov from the Gordon and Betty Moore Foundation was going to join us to give us a "research funders" perspective on the topic, but at the last minute was not able to make it.
- So we now have two speakers and one discussant for our symposium.
  - Ken Stanley is Professor of Computer Science at the University of Central Florida, a founding member of Uber AI Labs, and co-author of the book "Why Greatness Cannot be Planned."
  - Curtis Johnson is Manager of the Analytics and Cryptography Department at Sandia National Laboratories, where he studies information generation and spread as an evolutionary process in human and biological systems.
  - I'm Jeff Tsao. I'm a physical scientist and engineer by training, and am currently a strategic advisor to various US Department of Energy research funding offices with a strong interest in understanding how we can fund research more productively.
- I'll start with a 5m overview of the symposium.
- Ken will give a 25m talk on how creativity in artificial systems informs fostering human creativity.
- Curtis will lead 5m of Q&A on Ken's talk.
- I'll give a 10m talk on creativity as research's meta-objective.
- Curtis will lead 5m of Q&A on my talk.
- Then Curtis will lead 30m of general discussion on the overall symposium topic.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.

Intellectual Overview of Symposium

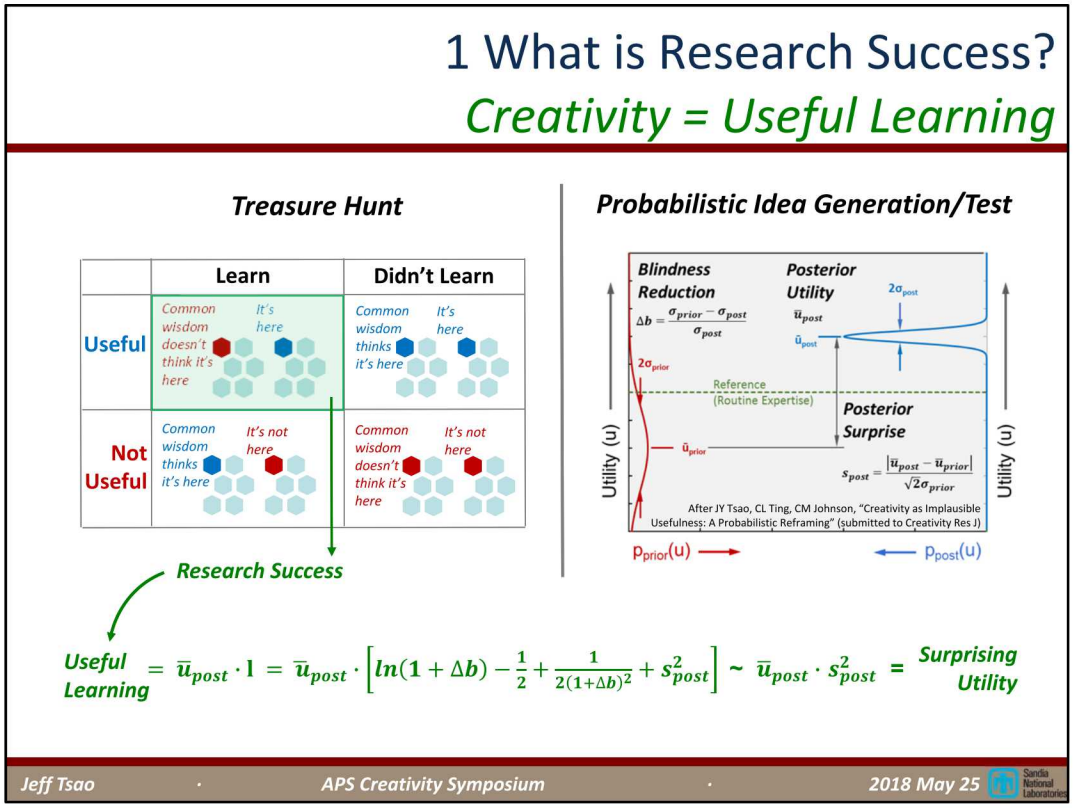


The extreme of exploratory research is both unpredictable *and* many levels removed from market feedback

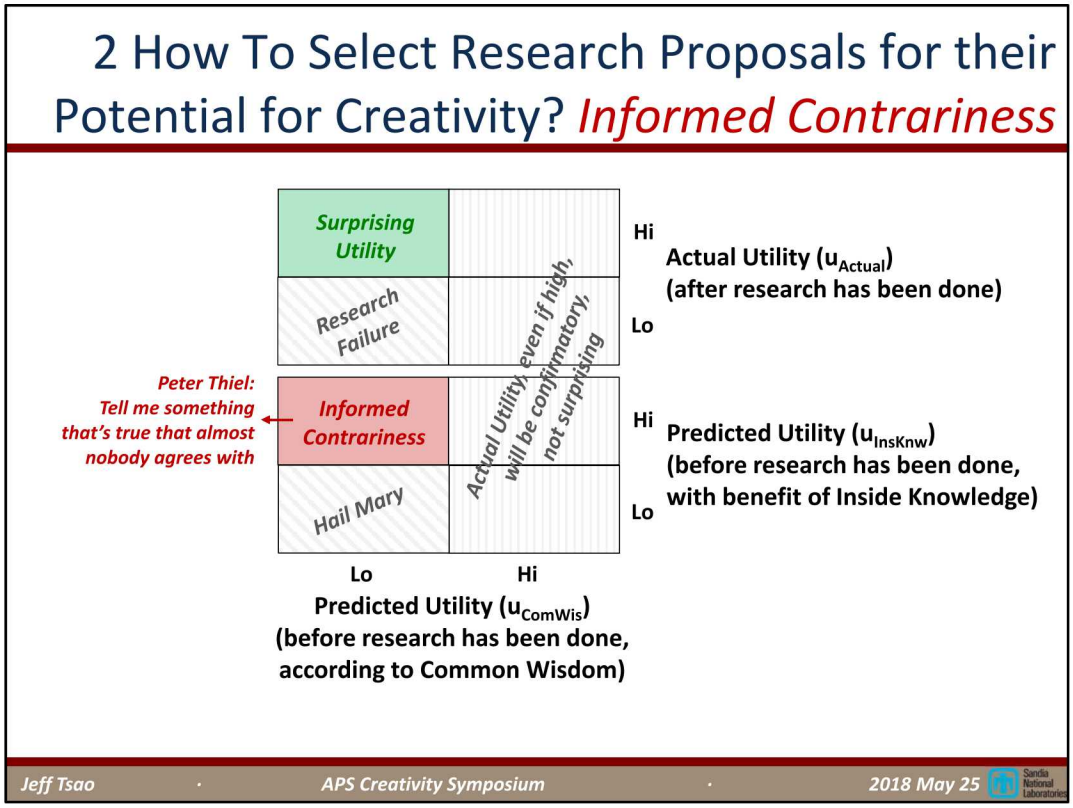
Current System:	**Assess usefulness and plausibility of pre-research objectives, then assess post-research accomplishments against pre-research objectives**
Problem:	**In exploratory research, post-research accomplishments can and should bear little relation to pre-research objectives**
Answer:	**Meta-objectives, not objectives! Novelty? Creativity?**

- Let me continue with an intellectual overview of the symposium.
- On the upper right of this slide I've drawn a box with two axes.
- The top axis, in red, is an exploit/explore axis. To the left you're exploiting the world, and what you are doing is much more predictable because you are making use of existing knowledge. To the right you're exploring the world, and what you are doing is much less predictable because you are creating new knowledge that you don't yet have. What this symposium aims to discuss is the extreme right side, where you can't predict what you will learn.
- The left axis, in blue, is a knowledge hierarchy axis. Knowledge of course is hierarchical – ideas and technologies build on other ideas and other technologies.
  - At the top, leaf-level of the hierarchy, ideas and technologies interact directly with consumers in the market, and the reward for the new idea or technology being innovative or not is high and fast – people either buy the iPhone or they don't, you either get rich or you don't. As a knowledge community, we don't need to judge whether to invest in a new idea or technology or not. Venture capitalists will make those judgements, and the quality of their judgements in turn will be judged by the market.
  - At the bottom, root-level of the hierarchy, ideas and technologies are many levels removed from consumers in the market, and the reward for the new idea or technology being innovative or not is low and slow. Einstein did not get wealthy because of special relativity, even though special relativity is one of the great intellectual revolutions of the last century, and one that the location sensor in every iPhone makes use of. It is this bottom part of the hierarchy that we are concerned with in this symposium. At the extreme of exploratory research many levels removed from the market, as a knowledge community we need to make our own judgements on what research to go forward with.
- How do we do it now?
  - Well, we have researchers write research proposals, we assess the usefulness and plausibility of the pre-research objectives in the proposal, then after the research is done we assess the post-research accomplishments against those pre-research objectives. Did you propose and then accomplish a new widget x with performance y?
  - The problem is that if you're really doing exploration, your uncertainty is extremely high, and your post-research accomplishments can and should bear little relation to your pre-research objectives. Einstein set out to understand why the speed of light was constant, but ended up discovering mass-energy equivalence – we don't want to penalize that, we want to celebrate it.
  - The tentative solution to this problem that we'd like to explore in this symposium is the following. It's not particular pre-research objectives that are important to assess. Instead, it is characteristics of those pre-research objectives that are important to assess – what we might call meta-objectives. Novelty and creativity, e.g., might be examples of such meta-objectives. If you're after novelty and creativity, then you can take credit for discovering mass-energy equivalence!
- With that, let me turn it over to Ken Stanley, who will give an artificial systems perspective on why assessing objectives doesn't work, why we need meta-objectives, and what some possible meta-objectives might be.

- TITLE. Hi again. We've just heard from Ken on the "myth of the objective": how targeting particular objectives in research can actually make ultimate progress towards research success less likely. Instead, we might target meta-objectives, e.g., novelty. In my talk I'd like to discuss another possible meta-objective: creativity itself.
- OUTLINE. Basically, I'll organize my talk around two questions.
  - Question 1 is: what should we consider successful research, after the research has been done? My answer will be creativity.
  - Question 2 is: how might we select research proposals before the research has been done, for their potential to lead to successful research after the research has been done? My answer will be informed contrariness.



- RESEARCH SUCCESS. Let's start with "What is research success?" As I said, my answer is creativity. And, by creativity I use Dean Simonton's and Mihaly Csikszentmihalyi's basic definition: useful learning. In other words, research is successful when you have learned something useful, and the more successful the research the more you have learned and the more useful what you have learned is.
- USEFULNESS-LEARNING MATRIX. This is probably reasonably intuitive, but its ramifications might not be so intuitive. To see the most important ramification, imagine the scenario on the left. You're a treasure hunter searching for a treasure that is in one of these six bins. There are four possibilities that I've sketched in this 2x2 matrix, organized according to whether the result of your search was useful or not useful, and whether you learned or didn't learn from your search.
  - In the lower right quadrant, common wisdom doesn't think the treasure is in this red bin, and after doing some research, you find that indeed it isn't in this red bin. This result is not useful because you still don't know where the treasure is, and on top of that you haven't learned because all you've done is confirm common wisdom, that it was not in the red bin.
  - In the lower left quadrant, common wisdom thinks the treasure is in this blue bin, but, after doing some research, you find that it isn't in that bin. This result isn't useful either because you still don't know where the treasure is, but at least you've learned something – common wisdom has been overturned.
  - In the upper right quadrant, common wisdom thinks the treasure is in this blue bin, and, after doing some research, you find that it *is* in that bin. The result is useful because you've found the treasure, but you haven't learned much because all you've done is confirmed common wisdom, which was that it *was* in that bin.
  - It is the upper left quadrant, shaded in green, where both usefulness and learning come together. Common wisdom doesn't think it is in the red bin, but, after doing some research, you find that it *is* in that bin. The result is useful because you've found the treasure, and on top of that you've learned – what you thought previously has been overturned. This is the useful learning, research success, quadrant.
- IDEA GST. I don't have time to go into the details, but let me just mention that one can quantify useful learning mathematically using an analysis like the one on the right of probabilistic idea generation and test. Your idea, before you've done the research, has some prior probability distribution over utility, sketched in red. Your idea, after you've done your research, has some posterior probability distribution over utility, sketched in blue. If research success is useful learning, then it might be something like the product of the mean utility of the posterior probability distribution, $\bar{u}_{post}$, times how much you've learned, l. How much you've learned is the so-called Kullback-Leibler divergence – an information-theoretic measure of how different the posterior and the prior probability distributions are from each other. And the Kullback-Leibler divergence, in turn, is this somewhat complicated expression. Again, I don't have time to go into the details, except to point out that this expression is dominated by this term s^2_{post}, which is a measure of how different the mean posterior utility is from the mean prior utility – that is, how surprised you were that the idea has the utility it turns out to have. In other words, useful learning is mathematically equivalent to *surprising* utility.
- USEFUL LEARNING = SURPRISING UTILITY. So this is what I meant by the ramifications being not so intuitive. We started with research success is creativity is useful learning, and we ended up with research success is surprising utility, where by surprising we mean the extent to which common wisdom is overturned.



- RESEARCH PROPOSAL SELECTION. OK, now let's move on to "How should research proposals be selected so as to maximize the potential for research success, or creativity?" For this we have to work backwards – inferring what an idea needs to look like *before* you've done the research, to optimize the odds of success *after* you've done the research.
- TWO MATRICES. To do this, I've drawn two matrices. The top matrix is the after-the-research-has-been-done matrix, and tells you whether your research has been successful or not. The bottom matrix is the before-the-research-has-been-done matrix, and tells you whether your proposal has the *possibility* of leading to successful research or not. Both matrices share the same horizontal axis: the predicted utility of a research idea before the research has been done, according to common wisdom. You can see that I've greyed out all of the bins for which this is high. The reason is that if common wisdom thinks the idea will have high utility, your research is *guaranteed* not to be successful. Even if the idea ends up having high utility, the utility will be confirmatory and not surprising, so you won't have learned. So you don't want to be in any of these four bins on the right. But what about these other four bins on the left?
- AFTER-RESEARCH MATRIX. Let's first look at the top after-the-research-has-done matrix. It has as its vertical axis the actual utility of a research idea after the research has been done. Research success is the green-shaded surprising-utility bin in the upper left. This is of course the desired outcome: the predicted utility according to common wisdom was low, but the actual utility turned out to be high. Note, though, that this is not the most *likely* outcome – that would be the research-failure bin in the lower left. After all, common wisdom is usually correct, that's why it's common wisdom. So if common wisdom predicts that utility will be low, most likely actual utility *will* be low.
- BEFORE-RESEARCH MATRIX. In other words, if you're going to go against common wisdom, you'd better have some good reasons for going against it. We might call these reasons "inside knowledge," something *you* know that no one else yet knows that makes you think you are right and common wisdom is wrong. This is illustrated in the bottom before-you've-done-the-research matrix. Its vertical axis is a predicted utility, just like that of the bottom axis, but with the benefit of this inside knowledge.
  - INFORMED CONTRARINESS. Within this matrix, it is the red-shaded informed-contrariness bin in the upper left which is where you want your research idea to be. You want common wisdom to think the idea will have *low* utility, but because of your inside knowledge you think the idea might actually have *high* utility. It's contrariness, because it goes against common wisdom, but it's an *informed* contrariness and so reduces the risk of going against common wisdom. This Isn't too different from what venture capitalists look for when they choose 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."
  - HAIL MARY. Of course, maybe your inside knowledge isn't very good, and if that's the case you'll be down in this lower left bin. You still might go through with the idea, if you are desperate and have no other options, but you'd be taking a Hail Mary.
- RESEARCH PROPOSAL SELECTION. So the essence of what we're looking for in a research proposal is informed contrariness. Informed contrariness doesn't guarantee success, you might still fail. But non-informed contrariness is worse, and not being contrary at all is also worse.

Summary and Questions for Debate!

- **Ken and Jeff's Common Theme**
  - **In selecting exploratory research proposals to fund:**
  - **shift from objectives (widget x with performance y)**
  - **to meta-objectives (leading indicators?) (novelty, informed contrariness, …)**
- **Ken and Jeff's Subtle Difference**
  - **Informed Contrariness $\subset$ Novelty:**
  - **Research proposal could be novel but not informed (perpetual motion machines that don't conserve energy)**
  - **Research proposal could be novel but not contrarian (why didn't *I* think of drinking coffee just before I nap)**

- OK, here's a high level summary.
- Ken and my common theme, I think, is that in selecting exploratory research proposals to fund, you can't base it on particular objectives, like widget x with performance y, because you'll miss too many things – like widget w with performance z. You have to base it on meta-objectives, like novelty or informed contrariness, that allow more freedom to find things that will ultimately be useful. But the freedom isn't unlimited. Just finding anything isn't useful – what we need are leading indicators for what will ultimately be useful.
- Where Ken and I differ is in exactly what those leading indicators should be. Ken's tentative leading indicator is novelty; my tentative leading indicator is informed contrariness. And note that informed contrariness is somewhat more restrictive than, or somewhat of a subset of, novelty.
  - A research proposal could be novel but not informed – as in a perpetual motion machine that doesn't conserve energy.
  - A research proposal could be novel but not contrarian – as in an idea that is new but is immediately obvious that it will work.
  - My question: is informed contrariness too restrictive?!