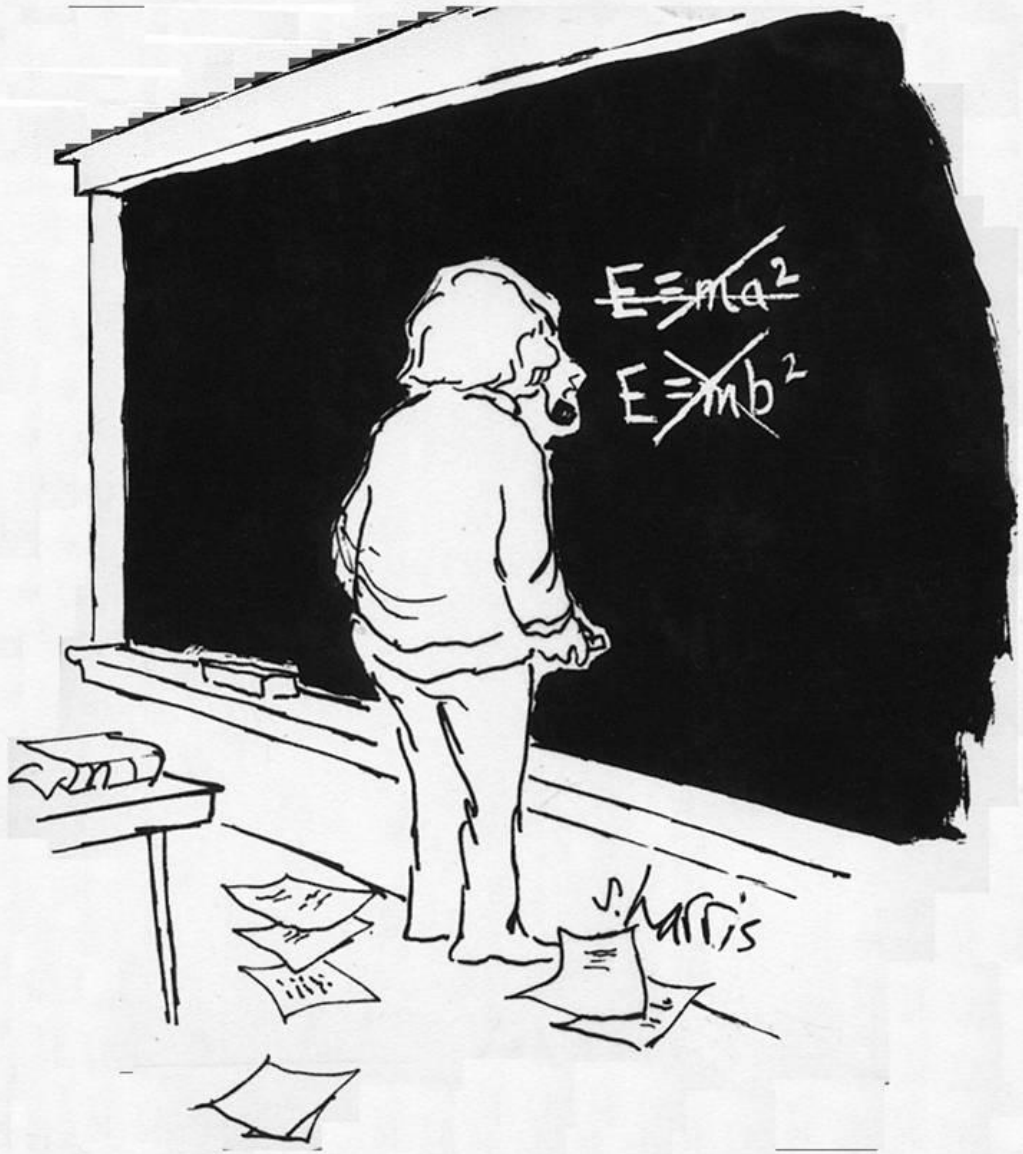

E=ma²
E=mb²
S. Harris

Scientific Creativity:

Discovery and Invention as Combinatorial

What's the basic thesis?

- Creativity, discovery, and invention all inevitably require some combinatorial process or procedure
- Although the same generic or superordinate mechanism applies to all creative domains, the specific content and characteristics of its operation will vary substantially by domain
  - e.g., scientific creativity will generally differ from artistic creativity
- Situation somewhat analogous to language: Despite the existence of thousands of distinct languages, linguistic universals or tendencies still exist with respect to acquisition, comprehension, and production

This thesis is nothing new!

- Introspective reports: e.g.,
  - Albert Einstein (1945) “combinatory play seems to be the essential feature in productive thought”
  - Henri Poincaré (1921) “ideas rose in crowds; I felt them collide until pairs interlocked, so to speak, making a stable combination”
- Psychological theories: e.g.,
  - Mednick (1962): “the creative thinking process as the forming of associative elements into new combinations which either meet specified requirements or are in some way useful. The more mutually remote the elements of the new combination, the more creative the process or solution” – hence the ubiquitous RAT!
- AI/Computer simulations: e.g., evolutionary algorithms (D. Simon, 2013)
- Creative products: e.g.,
  - Thagard’s (2012) analysis of 100 discoveries and 100 inventions of the highest order

Then what are my contributions here?

- First, I formalize the various types of combinations
  - Because not all combinations, nor even most, are creative
  - Because non-creative combinations may still lead to creative combinations
- Second, I specify the essential connection between combinatorial creativity and what has been variably styled “trial and error,” “generate and test,” “blind variation and selective retention,” etc.
- Third, I discuss the necessary properties of processes and/or procedures that generate potential creative combinations
- Fourth, I work out the implications for distinct domains of creativity, especially the contrasts between the sciences and the arts

Two constraints

- One, I focus on what’s going on in a single scientist's head, ignoring what also happens during lab meetings and other forms of group-level “brainstorming” (aka cognitive rather than social psychological)
- Two, I concentrate on problem solving, so that the interest largely concerns combinations that constitute solutions to given problems regarding natural phenomena, such as
  - Identification: What is it?
  - Explanation: How does it work?
  - Prediction: What happens if?
  - Invention: By what device can this be done?

Formalization

- Given a particular problem, ideational or behavioral combinations are generated that represent potential solutions, as in
  - **Galileo Galilei** applying his artistic training in chiaroscuro to recognize that the dark and light patterns on the lunar surface represented mountains
  - **Isaac Newton** integrating Galilean mechanics with Keplerian astronomy to produce his gravitational account of the solar system
  - **Charles Darwin** realizing that Malthusian population growth would cause a "struggle for existence" that drives evolution via natural selection
  - **Thomas Edison** combining his telegraph and microphone inventions, thus converting the transcription of Morse code into the recording of sound
  - **James Watson** tinkering around with cardboard models of the four nucleobases to discover the DNA coding (AT and GC sans Chargaff's rule)

Three parameters: *p*, *u*, and *v*

- At the instant that problem solving starts, those combinations may be described by the following three parameters (omitting subscripts):
  - *p* = the combination's *initial* probability or "response strength"
    - Where $0 \leq p \leq 1$ (i.e., not spontaneously generated during the initial session, to generated after a certain delay within the session, to instantaneously generated first)
  - *u* = the combination's *final* utility as a solution when fixed in the product
    - Where $0 \leq u \leq 1$ (i.e., completely useless, to merely satisficing, to satisfying all criteria)
    - N.B.: No possible temporal instability in the personal assessment of *final* utility
  - *v* = the *prior* knowledge of the combination's final utility before generation
    - Where $0 \leq v \leq 1$ (i.e., utterly ignorant, to informed hunch, to "justified true belief")
    - e.g., If 2 out 5 mandatory utility criteria are already known to be completely met, $v \approx .4$
- Hence a eightfold solution typology (cf. Tso, Ting, & Johnson, 2019)

Solution typology ("→" = "nears value of")

- Two expertise-driven solutions ($v \rightarrow 1$, hence nothing new is learned):
  - First, $p \rightarrow 1$, $u \rightarrow 1$, $v \rightarrow 1$
    - In words, the solution has a high initial probability because the individual already knows full well that it has a high utility
    - Hence, solution virtually automatic; representing *explicit* expert knowledge or skill
    - e.g., Where does this quadratic function maximize? Take the first derivative etc.
  - Second, $p \rightarrow 0$, $u \rightarrow 0$, $v \rightarrow 1$
    - In words, the would-be "solution" has a low initial probability because the individual already knows full well that it has a low utility
    - solution automatically rejected as a "non-starter" yielding "preselection"
    - rational suppression (aka extinction) representing *implicit* expert knowledge or skill
    - e.g., Don't bother making any assumptions that violate the laws of thermodynamics
    - Much mentoring devoted to "tacit knowledge" of this very nature

Solution typology ("→" = "nears value of")

- Two irrational "solutions" ($v \rightarrow 1$; maladaptive responses)
  - *First,* $p \rightarrow 1$, $u \rightarrow 0$, $v \rightarrow 1$
    - high probability even though the utility is already known to be low
    - irrational perseveration ("beating your head against a brick wall")
    - same solution repeated even though has never worked ("hope springs eternal")
      - "the definition of insanity is doing the same thing over and over and expecting different results" (falsely attributed to Einstein)
  - *Second,* $p \rightarrow 0$, $u \rightarrow 1$, $v \rightarrow 1$
    - Low probability even though the utility is already known to be high
    - irrational suppression
    - solution ignored even though known to work (presumably for a pathological motive)
  - Of course, scientists make neither of these mistakes (though artists might)

Solution typology ("→" = "nears value of")

- Prior ignorance of solution utility ($v \rightarrow 0$): Four possibilities
  - First, $p \rightarrow 1$, $u \rightarrow 1$, $v \rightarrow 0$
    - fortunate guess; like winning the lottery with your absolute favorite "lucky number"
    - predominant response just so happens to be correct, but prior knowledge remains zero
    - knowledge may be increased, but that knowledge is very often highly circumscribed
  - Second, $p \rightarrow 1$, $u \rightarrow 0$, $v \rightarrow 0$
    - highest probability response had no justification whatsoever, and is promptly dismissed
    - establishes the boundaries of one's expertise, as in **problem finding**, a critical event often underlying subsequent creative problem solving (e.g., Einstein & Infeld, 1938)
    - e.g., a failed prediction from a well-established theory (cf. Kuhn's 1970 "anomalies")
  - Third, $p \rightarrow 0$, $u \rightarrow 0$, $v \rightarrow 0$
    - mind wandering, daydreaming, environmental exploration, behavioral tinkering/play
    - low-probability thoughts and behaviors can make substantial contributions to creative problem solving whenever they "chance" upon a high utility solution ($u \rightarrow 1$)

Solution typology ("→" = "nears value of")

- Prior ignorance of solution utility ($v \rightarrow 0$): Four possibilities (cont.)
  - Fourth, $p \rightarrow 0$, $u \rightarrow 1$, $v \rightarrow 0$
    - In other words, high utility but low initial probability because the utility is poorly known
    - If $(1 - p)$ = *originality* and $(1 - v)$ = *surprise* (i.e., amount of new knowledge acquired),
    - then let creativity
    - $\boldsymbol{c = (1 - p)u(1 - v)}$, that is, the joint product of originality, utility, and surprise
      - cf. Boden (2004): novel, valuable, and surprising
      - cf. US Patent Office: new, useful, and nonobvious (with respect to "ordinary skill in the art")
    - Again, this represents a personal rather than a consensual estimate of solution creativity
    - Admittedly, the specified definition is just an "ansatz" (or "educated guess" $0 < v < 1$) that will prove useful to working out certain implications that test its utility
      - Specifically, the ansatz has six major repercussions that establish its high utility (i.e., $u \rightarrow 1$)

Six ansatz implications

- *First*, and most obviously, personal creativity *c* is a continuous variable that ranges from zero creativity to maximal creativity, with innumerable gradations between (to wit, $0 \leq c \leq 1$)
    - e.g., moderate creativity may be represented by $p = .5$, $u = .8$, and $v = .2$, which yields
        - $c = .32$, or about a third up the hypothetical scale
    - But less obviously, when $c << 1$, the creativity of a solution may represent an infinitely varied mixture of values for originality, utility, and surprise
        - e.g., the moderate creativity represented by $c = .32$ above could also result from the values $p = .2$, $u = .8$, and $v = .5$ or any of an indefinite number of other values
    - Yet the *qualitative* character of the creative solution will differ depending on which of the three criteria dominates
        - e.g., applied research might emphasize utility, but pure research might stress originality

Six ansatz implications

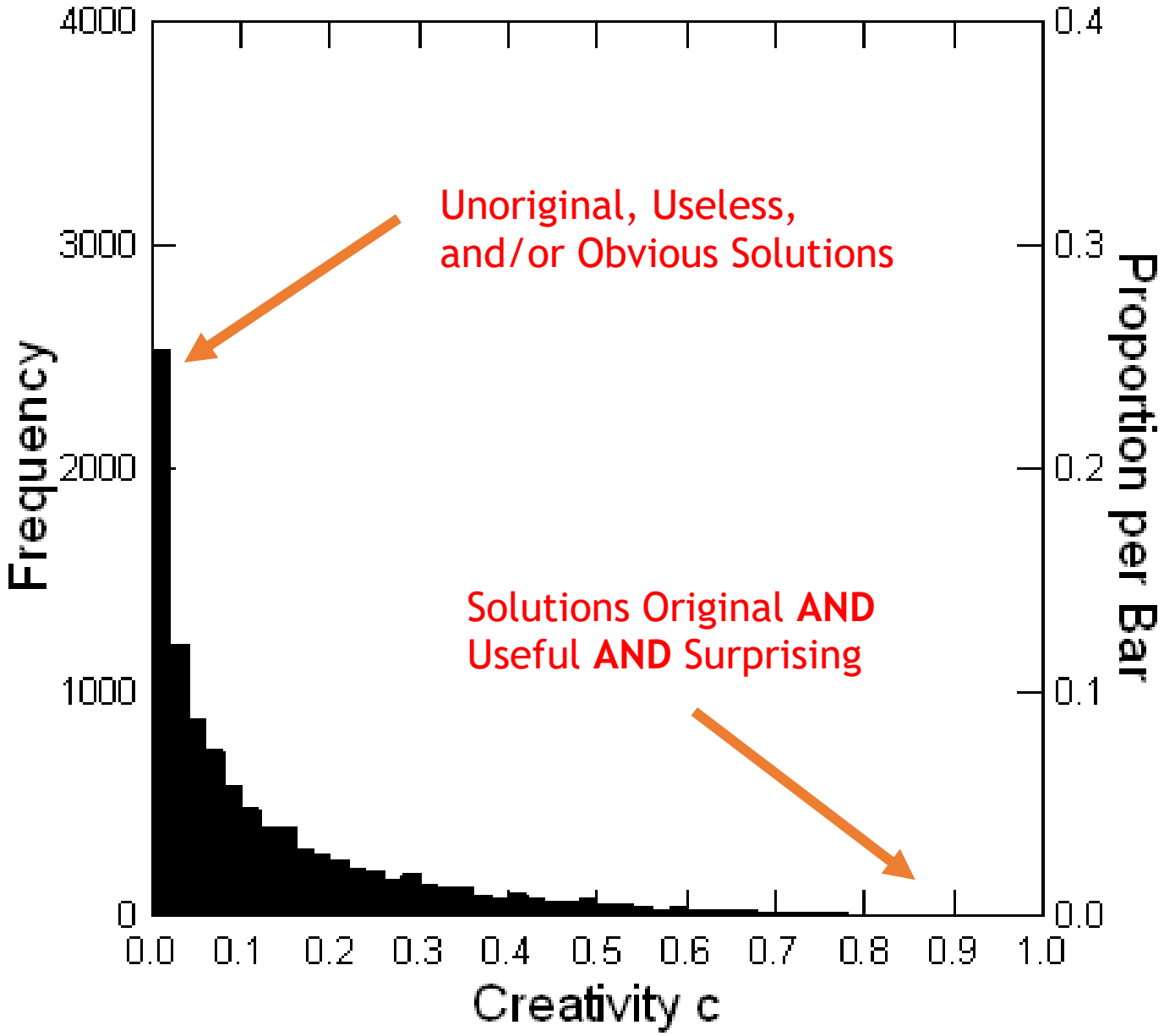
- *Second*, because the three criteria are combined via multiplicative rather than additive integration, each criterion becomes *necessary but not sufficient* for creativity (i.e., if any factor equals zero, then their product is zero)
  - e.g. bank safes made out of ordinary soap suds (p = .1, u = 0, v = 1 and c = 0)
  - cf. under additive integration, such as averaging originality, utility, and surprise, such putative "safes" would still be creative (viz. c = .30)
  - N.B.: In the arts, there may be instances where $u >> 0$ for suds safes
    - e.g., in a scifi movie where it turns out that the suds only *look* "ordinary" because they actually represent an alien force field that if touched immediately vaporizes the transgressor

Six ansatz implications

- *Third*, whenever $u > 0$ and $v < 1$, then c maximizes when $p = 0$
  - Hence, although an incubation period is not required to generate creative solutions, those solutions that require incubation because the initial probability is zero will tend to be more subjectively creative (the final utility and prior knowledge otherwise held constant at the specified inequalities)
    - e.g. Wallas's (1926) classic four stages of preparation, **incubation**, illumination, and verification (based on the experiences of superlative scientists, such as Helmholtz)
  - That said, creativity c is uncorrelated with incubation's temporal duration
    - because the latter is contingent on chance internal associations and external stimuli
      - e.g. "constrained stochasticity" (Carruthers, 2018; Simonton, 2003)
      - e.g. "opportunistic assimilation" (Seifert et al., 1995)
    - e.g. Archimedes's famous eureka experience would not have produced a more creative solution to the gold crown problem had he delayed taking a bath a day or more

Six ansatz implications

- *Fourth*, given multiplicative integration, then highly creative solutions must be far more rare than noncreative solutions (cf. additive integration, which yields a roughly normal distribution via the central limit theorem)
  - e.g., a simple Monte Carlo simulation in which values of the three parameters are randomly generated according to a normal distribution
  - N.B.: Same results with alternative distributions and with approximate constraints that delete irrational combinations (e.g., $uv \leq p \leq 1$)

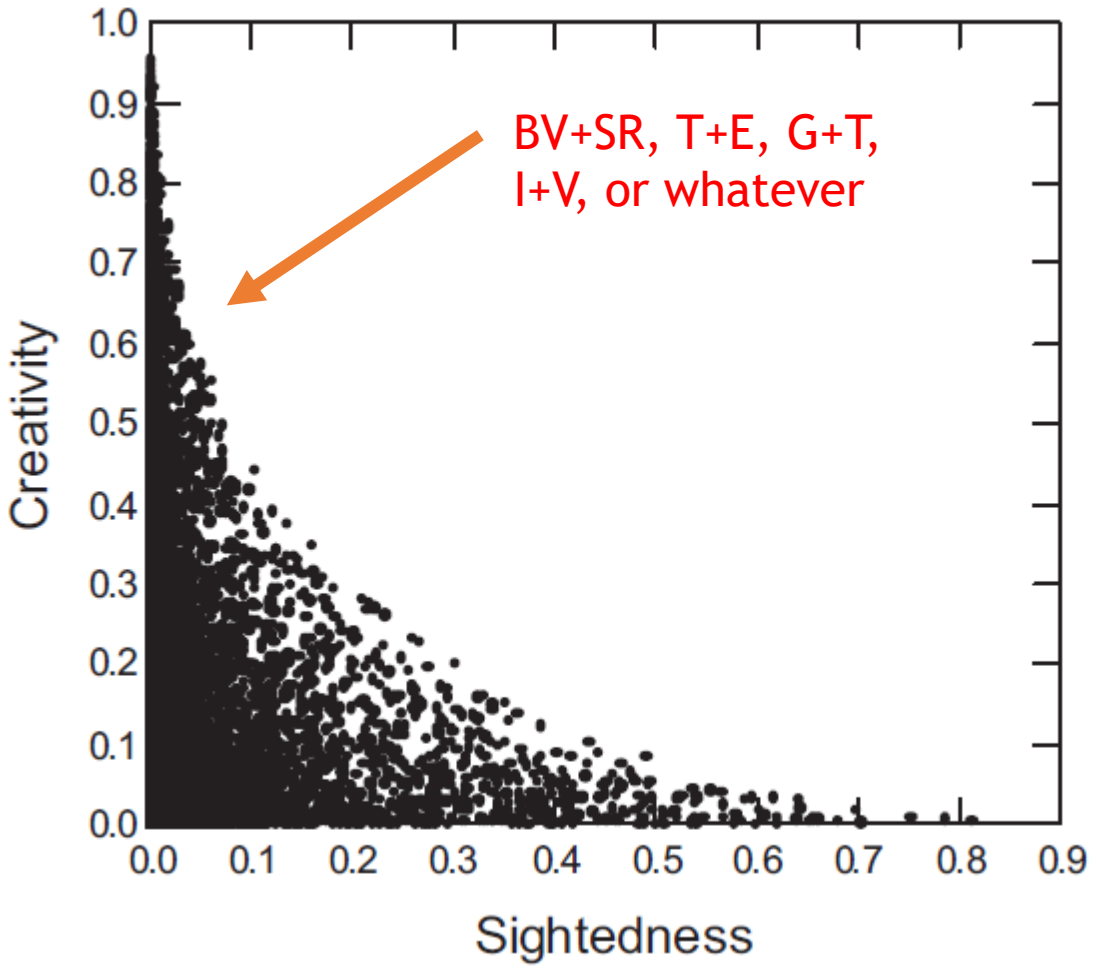

Unoriginal, Useless,
and/or Obvious Solutions
Solutions Original AND
Useful AND Surprising
Frequency
Proportion per Bar
Creativity c
4000
3000
2000
1000
0
0.4
0.3
0.2
0.1
0.0
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

Six ansatz implications

- *Fifth*, if the third criterion *v* is omitted, the resulting two-criterion "standard definition" makes neither logical nor psychological sense
  - The standard definition says creativity is (some unspecified) function of just originality and effectiveness (aka usefulness; Runco & Jaeger, 2012).
  - Yet if originality is defined as low probability, then why would a highly useful solution ever have a low probability?
  - The obvious answer is that the utility is hitherto unknown or incompletely so
  - Otherwise any rational creature would set the probability more proportional to the utility (viz., as $u \rightarrow 1$ and $v \rightarrow 1$ then $p \rightarrow 1$)
  - Indeed, without *v*, creativity is indistinguishable from irrational suppression!
  - Q.E.D. (this reductio ad absurdum inspired by reading Plato's dialogues)

Six ansatz implications

- *Sixth*, and last, because the solution utilities for creative combinations are unknown or incompletely known in advance of the generation of the potential solutions (i.e., $v << 1$), then those solutions must undergo a second step of directed evaluation or assessment
  - The two steps have been variously styled: **trial and error** (T+E; Bain, 1855), **illumination and verification** (I+V; Wallas, 1926), **generate and test** (G+T; various AI algorithms), **conjecture and refutation** (C+R; Popper, 1963), and **blind variation and selective retention** (BV+SR; Campbell, 1960)
  - 2nd step either external (Skinnerian) or internal (Popperian)(Dennett, 1995)
  - To illustrate, let us define a potential solution attribute called "sightedness" using the three parameters, namely, $s = puv$, where $s = 1$ when $p = u = v = 1$
    - Then another simple Monte Carlo simulation yields the following scattergram:



Albert Einstein “If we knew what we were doing, we wouldn't call it research” and
“Most of my intellectual offspring end up very young in the graveyard of disappointed hopes”

But where do the potential solutions originate in the first place?

- Answer: Whatever works! (cf. “anything goes” Feyerabend, 1975)
- Researchers have proposed an impressive number of combinatorial generators that can feed potential solutions into the selection hopper
- e.g., divergent thinking, remote association, cognitive disinhibition (defocused attention), primary (primordial) process (“regression in the service of the ego”), dreams, psychoactive drugs, organic brain disorders, synesthesia, intuition, overinclusive (allusive) cognition, mind wandering, analogy, conceptual reframing (frame shifting), broadening perspective, juggling induction and deduction, problem dissection, reversal, tinkering, play, heuristic and systematic searches, serendipity, Geneplore, Janusian, Homospatial, and Sep-Con Articulation thinking
- There’s **No Free Lunch**! *All* work *some* of the time, *none* works *all* of the time - and there’s no telling beforehand which will work best!

But where do the potential solutions originate in the first place?

- Even so, each and every generator shares one key characteristic:
  - the capacity to generate low probability potential solutions with unknown or incompletely known utility values
  - i.e., $p \rightarrow 0$ and $0 \leq u \leq 1$ while $v \rightarrow 0$
  - The latter ignorance then requires a utility evaluation or test
- Even highly inspired eureka moments, acceptable utility is by no means guaranteed: Hence, Wallas justified in adding the verification stage after the illumination stage (e.g., “Oh, shucks!” events)
- If the utility were already well known, then its generation probability would correspond proportionately, and the end result is explicit or implicit expertise, not creativity

But where do the potential solutions originate in the first place?

- Two phenomena illustrate the exceptional circumstances under which highly creative ideas are often generated in the sciences:
  - Internal circumstance: **Mind wandering** (Gable, Hopper, & Schooler, 2019).
    - "The bath, the bed, and the bus" (Boden, 2004, p. 25)
    - Because $p \rightarrow 0$ and $v \rightarrow 0$, then $0 \leq u \leq 1$ (viz. no preselection possible)
    - Hence, a creative combination can be spontaneously generated when engaged in some mundane or semi-alert activity, like taking a bath, waking up in bed, or boarding a bus
  - External circumstance: **Serendipity** (Cannon, 1940; Mach, 1896)
    - Classic examples: penicillin, electromagnetism, X-rays, phonograph, ozone, etc.
    - Note the unique parameter values: $p = v = 0$ *exactly*, but $u \rightarrow 1$
      - i.e., the highly useful combination could be neither spontaneously generated nor anticipated
    - Q: But what of Pasteur's "Chance favors the prepared mind"?
      - A: The scientist must have the requisite expertise to evaluate the combination's utility

How does scientific creativity differ from artistic creativity?

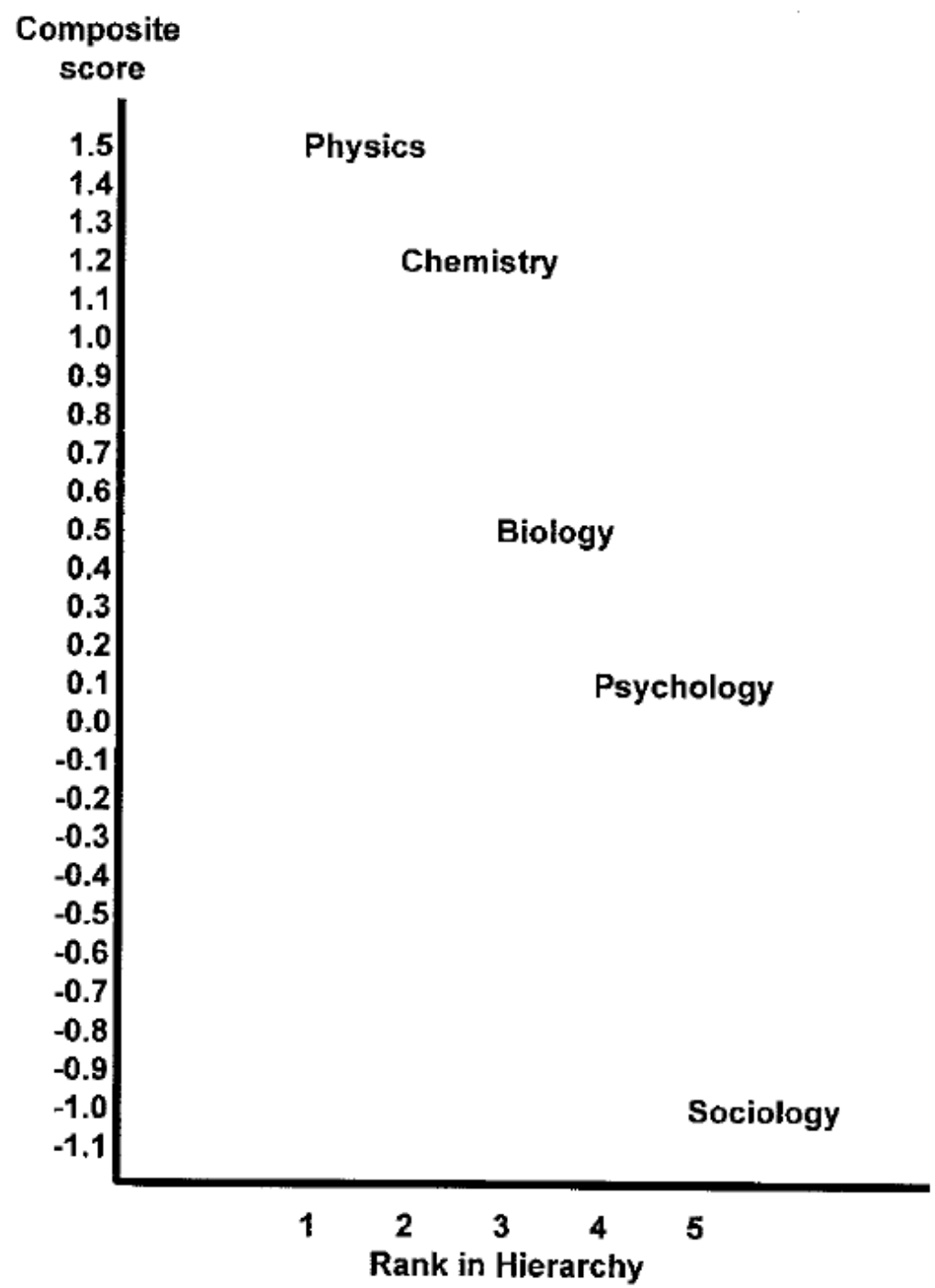
- Combinatorial processes and procedures permeate all forms of creativity, including in the arts
  - The latter is evident from multiple sources, such as introspective reports, theoretical models, computer simulations, single-case studies, etc.
    - e.g., quantitative analyses of the sketches that Pablo Picasso generated in route to *Guernica*: manifest variation-selection process that recombined images used in previous paintings and etchings (Damian & Simonton, 2011; Simonton, 2007; Weisberg, 2004)
    - e.g., computer programs that generate reasonable versions of human created music, art, and poetry (EMI, AARON, and RKCP; Cohen, 1990s; Cope, 2014; Kurzweil, 2000)
  - Yet fundamental contrasts distinguish scientific and artistic creativity
    - Many differences are obvious (e.g., background, training, personality, venues, audiences)
    - So here I concentrate the discussion on contrasts using the ansatz

How does scientific creativity differ from artistic creativity?

- First, scientific *combinations*
  - Prior knowledge value *v* more often higher (i.e., expertise driven hunches)
  - Utility *u* more precisely delineated (e.g., no “poetic license” allowed)
  - Initial probability *p* more likely higher (i.e., pure originality less valued)
- Second, scientific *products*
  - Larger proportion of the combinations collated into products represent scientific expertise rather than creativity (i.e., $p = u = v = 1$)
    - lit reviews, methods and statistics boilerplate, routine citations, standard format, etc.
- Third, scientific *consensus*
  - Personally assessed “little-c” creativity more strongly corresponds to consensually assessed “Big-C” creativity [i.e. $(p, u, v) \approx (P, U, V)$]

Postscript

- The foregoing contrast between scientific and artistic creativity represents only the first cut
  - On the one hand, artistic creativity can vary from the highly formal domains like architecture to the highly expressive domains like poetry
    - with corresponding repercussions (cf. Ludwig, 1995)
  - On the other hand, scientific creativity can also vary from highly paradigmatic "hard" sciences to the far less paradigmatic "soft" sciences (Fanelli, 2010; Fanelli & Glänzel, 2013; Simonton, 2004, 2015)
    - to wit …


Composite score
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1.0
-1.1
Physics
Chemistry
Biology
Psychology
Sociology
1
2
3
4
5
Rank in Hierarchy