"*Constructive Skepticism*" Volume 3 – Notebook #I: *Model Risk* **– "Spinach"**

This series of workbooks, handbooks, and notebooks collects thoughts, definitions, and sources focused on using "*Constructive Skept*icism" to make good individual decisions. As you can see from the title, and the first sentence, words with a specific "*Meaning*" anchored the work of a referenced author show up in "*Italicized Quotes*" to alert the reader that they have a specific "*Terms-of-Art*" definition in the Glossary, and to remove the confusion that may come from reading through the text with an implicit, and different understanding.

Having witnessed the difficult retirements of two prior generations, first as a child, then as an adult, the topic of retirement planning became a personal focus of attention. This experience with prior generations showed that retirement looked like a multifaceted problem worth trying to solve over a lifetime of work. Not only would the result apply to this writer, but they would also apply to nearly all readers.

This lifetime of work showed that retirement planning starts as a decision problem based on hard-to-see clinical inputs, and that retirement management continues as a measurement problem based on observable outcomes. A retirement plan provides a long-term, adaptive, and protective structure for the client, a structure that can protect their health, wealth, and standard of living. This means that a retirement plan starts in the domain of moral choice, addressing values, goals, and individual ambiguity before continuing in the domain of economics, addressing issues about asset types, income flows, and historical averages.

This personal focus resulted in the creation, and the selling of several start-ups, as well as the development of "*Tools, Checklists & Processes*", including the analysis of the household balance sheet, the calculation of a dollar measure for risk capacity, and retirement planning recommendations based on risk allocations. CTRI, a membership-based, not-for-profit R&D institute, extends and expands these theories & best practices beyond what was developed in prior ventures.

A quote from the book "*Old School*" written by **Tobias Wolff**, and first published in 2004 illustrates the purpose of this personal focus and continuing work: "*In former time Arch had supposed that his sense of being a distinctive and valuable man proceeded from his own qualities, and that they would sustain him in that confidence, wherever he happened to be. He's never imagined that this surety was conferred on him by others, by their knowing and cherishing him. But so it was. Unrecognized he had become a ghost, even to himself.*" This quote reveals a key aspect of this quest to make good individual, business, and investment decisions in the context of retirement. It reveals the importance of sharing what we know with others for the mutual benefit of all involved. "*Meaning*" comes from what we do for others, and what they do for us.

This effort revealed the dangers and opportunities that arise when deriving "*Meaning*" from cooperation turns into deriving "*Meaning*" from consensus. Pressed for time and resources in real-life, we cannot double-check everything, and we must trust what

eventually become much too much. This reality of living in a limited world, and reliance on others opens the door for "*Willful Ignorance, Error & Deceit*".

This combination of cooperation, consensus, danger, and opportunities affects daily life as well as scientific research. Thus, it pays to find ways to limit the danger, and take advantage of the opportunities, especially in fields like retirement planning where dangers compound over time, opportunities take time to become fruitful, and the time to make corrections becomes shorter and shorter as we get older.

The workbooks collected in Volume 1 of this series of books, and the handbooks collected in Volume 2 develop "*Tools, Checklists & Processes*" to help readers find the "*Meaning*" of research papers, and books. Written in 2022 and 2023 using a systematic short-form based on writing a two-page section each and every day, they will remain available on Substack until their eventual print publication on Amazon, and include:

Volume 1: A Book of Connections in Four Workbooks
- Workbook #I: Our Shared Humanity
- Workbook #II: Making Good Individual Decisions
- Workbook #III: Making Good Business Decisions
- Workbook #IV: Making Good Investment Decisions

Volume 2: A Book of Collections in Four Handbooks
- Handbook #I: Glossary List and "Terms-of-Art" Definitions
- Handbook #II: Author Profiles
- Handbook #III: References
- Handbook #IV: The Template for Reading Research Papers

Publication on Amazon begins in 2024 with the print version of Template for Reading Research Papers so that readers can start with the end-product in hand, and use this collection of "*Tools, Checklists & Processes*" to "*See for Yourself*". The ideas developed in these notebooks, written in long-form with no page limitation, and collected in Volume 3 will also become available on Substack until their eventual publication in print on Amazon.

These ideas gathered in these volumes rest on "*Axioms, Assumptions & Hypotheses*" described in the workbooks from Volume 1, and further defined in the handbooks from Volume 2. These "*Axioms, Assumptions & Hypotheses*" start with the following "*Observations*":

- "*Brains*" exists to manage "*Motions*" through "*Predictions*" (See Author Profile about **Rodolfo Llinas** in Volume 2), and continue with
- "*The Map is not the Territory*" (See Author Profile about **Alfred Korzybski** in Volume 2).

This reading note - Volume 3: Notebook #I: Model Risk – uses the Template for Reading Research Papers to find the shared "*Meaning*" of "*Observations*", "*Perceptions*",

"*Predictions*", and "*Motions*" about model risk, one key researcher at a time, and starting with a review of **Mike Sutton**'s work.

Mike Sutton

The process of writing these workbooks, handbooks & notebooks started with reading a 2010 paper titled "*SPINACH, IRON and POPEYE: Ironic lessons from biochemistry and history on the importance of healthy eating, healthy skepticism and adequate citation*", and written by **Mike Sutton**, a former Reader in Criminology at Nottingham Trent University, UK who now writes books, and maintains a blog, www.dysology.org .

Sutton's paper hit close to home, as I remembered my mother making me eat spinach, that I did not like to eat, because she believed it was high in iron. Given number of follow-up papers and blog posts written by other authors since 2010, **Sutton's** paper must have hit close to home with other people as well. The word "*Spinach*" became a personal code-word for things we think unquestionably true but look ambiguously false after asking a few questions.

"*Spinach*" led to a personal interest in checking out received wisdom by reading foundational research papers focused on retirement planning. **Sutton's** paper, and the controversies that it generated show that the objects that populate the "*Mind-maps*" that we use to make "*Predictions*" in order to direct "*Motions*" in our "*Task Environments*" come from "*Observations*" filtered by "*Media*" into "*Perceptions*" of reality, and raise the following question: How much self-awareness of this filtering of "*Perceptions* do we need in order to make good decisions about retirement planning?

Eventually, this personal interest took an institutional form with the foundation of CTRI, and the scope of this interest in "*Spinach*" expanded beyond retirement planning to include general questions about making good individual, business & investment decisions. These questions drive the topics for Volume 3: A Book of Illustrations with a Growing Number of Notebooks.

The myth of "*Spinach & Iron*" provides an opportunity explore a mild controversy with a sufficient level of human interest and published detail in order to create an analogy that we can use with wilder controversies. This analogy proves rich in useful features for making good decisions from reading research papers:

- The "*Perspective*" of the researcher works like a light in darkness. To see something, you must shine a light on it, and the color of light from one researcher differs from another, up to the point of changing the nature of what we can see.
  - o What did Sutton see?
- The "*Domains of Knowledge*" mastered and used by the research anchors, and limits their scope, much like the location of a lamp post in physical space limits the area, and volume that its bulb can illuminate.
  - o What did Sutton know?

- The "*Historical Lineage*" of the researcher anchors the research in time, opening the door for the "*Real Story*" from individual discoveries, as well as the "*Good Stories*" from lost group memories.
    o When did Sutton work?
- The "*Purpose*" of the researcher impacts the reader based on their position in the a group, a "Dominance Hierarchy". Some seek to maintain the consensus, the status-quo, others stand outside-looking-in and seek change and novelty.
    o What did Sutton want to do?
- Study design that researchers can use range from weak "*Methodologies*" to strong "*Methodologies*". This choice has a structural impact on the "*Meaning*" of their work.
    o How did Sutton do his work?
- Researchers can measure "*Effects*" with a range of quantitative "*Methods*" whose mathematics come from abstractions applied to empirical problems with various degrees of fit, and fitness for use.
    o What did Sutton measure?
- The abstractions used by researchers work like a scaffolding built upon specific "*Axioms, Assumptions & Hypotheses*" that create a "*Small World*" whose features may or may not remain valid when applied to empirical problems in the "*Large World*", impacting the "*Meaning*" of the research when they do not apply.
    o What did Sutton believe?
- Validating the "Meaning" of research starts with understanding the "*Statistical Meaning*" of its measurements. Empirical measurements come with structural constraints and limitations.
    o What did Sutton test?
- Validating the "*Practical Meaning*" of research continues with understanding what it changes in the reader's "*Perceptions*", and ability to make "*Predictions*".
    o What did Sutton change?

According to Wikipedia [in full awareness of Sutton's warnings about the reliability of its content], following graduation with a Bachelor of Arts in Law from the University of Central Lancashire in 1983, **Michael Robert (Mike) Sutton** (1959 - Present) had a career focused on the statistical and research aspects of criminology. After the Dot-com Boom, he focused his attention on internet-based research, and developed a method he called Big Data Internet Dating in order to establish the veracity of what he called myths ranging from "*Urban Legends*" (such as "*Spinach & Iron*") to claims of scientific discoveries (such as **Darwin's** theory of Natural Selection). For instance, he self-published a book in 2014, titled "*Nullius in Verba: Darwin's Greatest Secret*" seeking to show that **Charles Darwin & Alfred Wallace** copied the theory of Natural Selection from a book published by **Patrick Matthew** in 1831, and titled "*On Naval Timber and Arboriculture*". In his 2010 paper, **Sutton** points out that "*Nullius in Verba*", the motto of The Royal Society, means "*On the Word of No One*", or "*take nobody's word for it*". **Sutton** champions skeptical inquiry, and "*checking the research behind assertions of fact*".

Before **Sutton's** paper, the "*Good Story*" of "*Spinach, **Popeye**, and Iron*" looked like an understandable story that made so much sense that it became an academic urban legend

that Sutton himself repeated in a presentation in 2009. The story has the structure of a good morality play, emphasizing the importance of doing good research, disbelieving ancient authorities, and being mindful of false positives. You may even have heard it yourself, used in conference presentations and sales pitches as one of the "*Tools*" of "*Fear, Uncertainty, & Doubt*" to try to "*Nudge*" your perspective, and to lead you to agreeing with them.

In his 34-page paper from 2010, **Sutton** describes his own involvement with the story of as follows:

- "My own involvement in this Spinach Popeye, Iron Decimal Story (SPIDES) began when I told it, at Manchester University, to an audience of academics, criminal justice professionals and civil servants during my introduction to an academic paper on the impact of bad data on policy making (Sutton and Tseloni 2009). The precise version of SPIDES that I told was taken for the most part from "Spinach – The Truth" (BBC 2006): "*So what about Popeye, then? During the 1930s Popeye, probably the worlds' most famous consumer of spinach, was indeed credited with a 33% increase in the consumption of spinach in the USA. These days the brand Popeye of spinach is one of the market leaders. The mythical strength-giving properties of spinach are, however, mostly credited to a simple mistake concerning the iron content of the vegetable. In 1870, Dr E von Wolf published figures which were accepted until the 1930s, when they were rechecked. This revealed that a decimal point has been placed wrongly and that the real figure was only one tenth of Dr von Wolf's claim.*"

When **Sutton** decided to verify the assertions made in this "*Good Story*" he sought to answer the following questions:

- "*Who is von Wolff, and where are his erroneous findings recorded? And where exactly is the evidence that any other scientists, such as von Bunge, and others up until the mid-1930s misplaced a decimal point in their presentation of findings regarding the iron content of spinach?*"
- "*Who were the 1930s chemists that discovered the decimal place error and where are their findings recorded?*"
- "*Since a correlation is no proof of causation, evidence is required to demonstrate that Popeye's creator Segar was indeed misled by erroneous science about iron in his choice of spinach for Popeye.*"
- "*Where is the evidence to support the claim that Popeye was alone responsible for increasing US spinach consumption by 33 per cent – and that spinach consumption had indeed increased in this way in the USA between the late 1920's and early 1940s?*"

These four questions reflect **Sutton's** "*Perspective*", and show that he worked to extend his "*Domain of Knowledge*" from criminology to biochemistry, nutrition, and public health. He is looking for a culprit, a person that can be blamed for starting the confusion. On the other hand, one could also choose questions that reflect a business "*Perspective*" and "*Domain of Knowledge*" to get to a different bottom-line, and such that questions about **Popeye** present a decision problem - Why eat spinach? - and questions about iron

present a measurement problem - What is the iron content in mg of iron per 100g of spinach? The measurement questions should be reproducible without the need for detective work: Is it (i) 340 mg of iron per 100g of dry spinach, (ii) an average value between 20.7 mg and 53 mg per 100g with a frequently mentioned value of 34 mg per 100g, or (iii) a value around 2.2 mg or iron per 100g of fresh spinach?

Sutton explains his intent and summarizes his finding in the abstract of his 34-page paper as follows: "*To inform knowledge in research methods and dissemination ethics for the natural and social sciences, this article reinforces the importance of citation to support all assertions of fact. New findings are presented for the history of biochemistry, nutrition, psychology, medicine, and the social sciences. Bio-chemistry papers and scientific news reports from the 1930's seriously undermine a long standing truism that in the 1920s and 30s, bio-chemists, nutrition experts, public health policy makers, and E. Segar the creator of the newspaper comic strip Popeye were misled either by a decimal place error in 19th Century published research, or else by erroneous interpretation of 19th Century scientific findings, to exaggerate the iron content of spinach tenfold. Further, the failure to study original sources is evidenced in a multitude of completely erroneous publications claiming that they apocryphal errors cause Segar to choose spinach for Popeye's super human strength. In fact, Segar chose and promoted spinach for its vitamin A content alone*."

While **Sutton's** paper confirms that according to current benchmarks spinach is lower in iron than the earliest estimates and by a factor of ten, thus solving the measurement problem, his place in time (the "*Historical Lineage*" of the Dot.com Boom), and choice of "*Methodology*" (Google Search) turned the once understandable & memorable story into a confusing story – muddying the clarity of solving the measurement problem.

Sutton is a criminologist working to validate the reproducibility of research results, and his own style and motivations become part of the story. He ends up obsessing, like a real life **Columbo**, to prove the existence of an intentional lie to cover-up a famous British researcher's mistaken and careless assertion from authority, and without accurate citation. The arc of the "*Good Story*" goes upside down. Some of the German scientists may, or may not exist. The good guy in the "*Good Story*" may, or may not be the bad guy in the "*Real Story*". The cover-up may have started somewhere else, and confused possible "*Perceptions* beyond repair. The required level of research becomes mind-numbing: One would have read all of **Popeye** newspaper cartoons, and in ascending order of publication to perhaps answer one of four key questions. There may be other important questions to ask, and researching these other questions, may lead in other directions about the negative, or not so negative, perhaps even positive impact of oxalic acid on iron absorption during digestion, and depending upon other diet considerations.

Sutton's observed that many books and papers, above and beyond the issue of "*Spinach & Iron*" provided, at best, inaccurate citations. Ironically, blog author **J. F. Derry** (https://mrsuttonntu.wordpress.com/) accuses **Sutton** of committing similar errors. Additionally, **Sutton** complains, in updates on this website, that the generally supportive authors of a 2019 paper (**Michael Mielewczik & Janine Moll**) titled "*Spinach in

Blunderland: How the myth that spinach is rich in iron became an urban academic legend" are doing the same thing toward him. Fictional, erroneous, or inaccurate citations remain an interesting problem, but these mostly ad-hominen arguments suggests that continued progress requires that we find other research papers besides **Sutton's**.

Thomas Foster

This "*Spinach*" adventure shows that "*Beliefs, Knowledge & Wonder*", the three levels of quality for the objects that populate our "*Mind-Maps*", come with various levels of individual ambiguity and reproducible validity. "*Spinach*" also shows that efforts to resolve this ambiguity, and to improve validity do not guarantee certainty. This creates model risk.

Usually, the analysis of model risk focuses on "*Methodology*", "*Methods*", and "*Statistical Meaning*" of the paper. However, looking at **Sutton's** paper suggest that we also need to look at the "*Perspective*", "*Domain of Knowledge*", "*Historical Lineage*", and "*Purpose*" of the author, as well as the "*Practical Meaning*" of the research. The new, top question becomes: How far must we take this validation process in order to make good decisions?

Over the last five years, the repeated "*Observation*" of "*Spinach*" in research papers led to the incremental development of the Template for Reading Research Papers. This template, based on the work presented in the workbooks from Volume 1, and the handbooks from Volume 2, summarizes the "*Tools, Checklists & Processes*" that you, as a reader can use to "*See for Yourself"* in order to trust your own "*Perceptions*", justify the accuracy of your own "*Predictions*", and make your own good decisions.

The iterative development of the Template started with **Thomas C. Foster**'s framework for reading literature because – as written by **Tobias Wolff** in "*Old School*" the literate appeared to know "*what was worth knowing*". **Foster's** 2003 book "*How to Read Literature Like a Professor*" presents the recurring themes that appear in novels. His framework reveals how to look, to see, and to read the symbolic significance of stories. He explains literature's "*Reference Narratives*" and uses them like "*Heuristics*" to find the practical "*Meaning*" of texts.

Foster offers a "*Checklist*" that maps easily into **John Boyd's** "*OODA Loop*" (See matching Author Profile in Volume 2), and **Charles Darwin**'s theory of Natural Selection:

- **Why**: The "*Real Story*" for a quest is always self-knowledge, an evolutionary story of "*Differentiation*".
  - o The dangerous situation, the crisis creates the emotional appeal around the specifics of the "*Good Story*".

- **Who**: People eat or drink together, as a form of communion: Would you accept a dinner invitation from someone you do not care for (an evolutionary story of "*Selection*")?
    o "*Observe*" with the eyes of the hero to provide the proof statement of the situation
- **What**: Power & Control: Ghosts, villains, and vampire illustrate a refusal to respect the autonomy of other people.
    o "*Orient*" the "*Good Story*" around the villain to assert an opinion, and to evaluate a plan of action, an evolutionary story of "*Change*".
- **How**: How much will the protagonists give up (e.g. their soul) in order to satiate their deep, motivating hunger?
    o "*Decide*" and (in business terms) make thoughtful requests to compound luck through the mastery of comparative advantage, an evolutionary story of "*Amplification*".
- **Where/When**: Don't be the "*Red Shirt*" (Star Trek fans know this "*Reference Narrative*"): Never stand next to the Hero. Negative emotions cannot be dispelled, only redirected against other characters, an evolutionary story of "*Dominance*".
    o Achieve transcendence through the cumulative effect of completed actions, "*Motions*", and resolved promises.

Foster shows that making good individual decisions in the arc of the life of a literary character follows a "*Checklist*" of patterns that use predictable "*Tools*", and repeatable "*Processes*" where each decision comes to a conclusion in specific windows of time. Asking Why, Who, What, How and When/Where proved useful, but insufficient to tease out the "*Statistical Meaning*" and the "*Practical Meaning*" of research papers.

The Template for Reading Research Papers

After years of incremental developments that blended the quantitative with the qualitative, the Template for Reading Research Papers grew to sixteen pages. It starts with a front page of instructions about the "*Process*". It continues with a one-page summary of its eight steps, including key sample questions. Finally, it closes with fourteen pages of details that include charts, tables and models in support of deeper sample questions for each one of its eight levels of analysis. The three-step "*Process*" for reading research papers includes:

- Document the citation, & jotting first impressions using the one-page summary,
- Take detailed reading notes using:
    o The sample questions provided with the charts, tables & models,
    o Custom questions developed by the reader from looking at the Template's charts, tables & models, or
    o Custom questions developed by reader based on reading the details, connections & references provided in the other workbooks
- Reproduce the research quantitative models in order to replicate its results.

Instructions for using the Template with each one of these three-steps include:

- First Step: Scan the Abstract, Introduction, Headings, Conclusion, & References to understand the structure of the paper
    o Using the one-page summary, does this first step provide sufficient information to answer questions about "*Perspective*", "*Context*", "*Purpose*", and "*Historical Lineage*"?
- Second Step: Take Notes for all levels of analysis in the Template
    o Using the one-page summary, does this confirm the information gathered from the first step, and provides sufficient information to answer questions about "*Methodology*", "*Methods*", "*Axioms, Assumptions & Hypotheses*", and "*Meaning*"?
- Third Step: Dig deeper
    o Using the Sample Questions in the template, does this change first impressions?
    o Using custom questions, does this change prior impressions?
    o Replicating the results directly, or indirectly with third-party meta-analyses, does this confirm the validity of the meaning of the paper?

The one-page suummary of sample questions for the first reading step include:

1- The "*Perspective*" of the author(s) of the paper
    a. Does this research have a (i) Descriptive ("*What if I See?*" questions), (ii) Prescriptive ("*What If I Do?*" questions), (iii) Predictive ("*Why*?" questions), (iv) Pragmatic (Individual Clinical Ambiguity), or (v) Agenda-driven perspective?
    b. Does this author have the perspective of an (i) Analyst, (ii) Clinician, (iii) Modeler, (iv) Statistician, (v) Synthesist, or a (vi) Polemicist?
    c. Does this author belong to a specific academic or business ecosystem?

2- The "*Domain of Knowledge*" for the author & the paper
    a. What Domain or Domains of Knowledge characterize this research?

3- The domain-specific, "*Historical Lineage*" of the author & the paper
    a. What antecedent, as well as descendent schools of thought & authors provide the historical lineage for this research?
    b. What historical "*Metaphors*" (e.g. stimulus-response, machinery mechanism, computer-like cognition, etc.) anchor the thinking in this paper?

4- The "*Purpose*" of the paper
    a. How would you describe the intended audience(s) for this research?
    b. What does this research seek to amplify (e.g. agreement) or suppress (e.g. judgement) for each audience?

5- The "*Methodology*" of the research

      a. What analytical programs, and related "*Maintenance Programs*" or "*Repair Programs*" does this research use?

      b. Does the selected research methodology match the intended audience(s) / "*Task Environment(s)*"?

6- The "*Methods*" used in the research

      a. Does this research engage in the application of "*Statistical Rituals*" such as "p-values", or in the application of insightful "*Statistical Thinking*"?

7- The foundational "*Axioms, Assumptions & Hypotheses*" that support the research

      a. How do this research's "*Axioms*, *Assumptions & Hypotheses*" limit its validity?

8- The "*Meaning*" of the paper as a valid reference

      a. How would you summarize the (i) Relevance, (ii) Coherence, (iii) Data Quality, (iv) Certainty, (v) Practical Significance, (vi) Accuracy, (vii) Precision, (viii) Statistical Significance, and (ix) Interpretation of this research for its intended audience(s)?

      b. What "*Named Items*" does this research (i) Intensify, (ii) Displaces, (iii) Makes Obsolete, and (iv) Reverses at Scale?

      c. Decide if you want to (i) Read the paper in details, and (ii) Replicate the results.

At this point, readers have a set of "*Tools, Checklists & Processes*" to manage Model Risk. They can use the Template for Reading Research Papers to find and the "*Statistical Meaning*" (to solve a measurement problem), as well as the "*Practical Meaning*" (to solve a decision problem) of a research paper.

This notebook will continue to explore Model Risk with future posts focused on the following authors and ideas:

- Richard Prum: "*Small Worlds*" Models and Natural Selection

- Mandelbrot: "*Small Worlds*" Models and Randomness

- Marshall McLuhan: Fooled by the "*Merchants of Attention*"

- Carlo Cipolla: Stupid, Helpless, Bandit, or Intelligent researchers?

- Ken Thompson: There can be no trust in machine-conjured reality

- Gerd Gigerenzer: Keep it simple in the "*Large World*"