



# ADVANCED ANALYTIC RIGOUR

**Ashley Barnett & Tamar Primoratz**  
The Hunt Laboratory for Intelligence Research  
The University of Melbourne



October 2020

# 1. Introduction



# Today's seminar:

## Fundamentals of Evaluating Reasoning



- Course objectives
- Training Method
- Reasoning structure
- Argument Mapping
- Evaluating Arguments

Jack is looking at Anne, but Anne is looking at George. Jack is married, but George is not. Is a married person looking at an unmarried person?

**A: Yes**

**B: No**

**C: Cannot be determined**



In a lake there is a patch of lily pads.  
Every day, the patch doubles in size.  
If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?



You are getting tested for a virus that affects 10% of the population.

This is what we know about the accuracy of the test:

- If someone has the virus, the result will be positive 80% of the time and negative 20% of the time.
- If someone doesn't have the virus, the result will be negative 80% of the time and positive 20% of the time.

Your test comes back positive.

What is the probability you have the virus? Pick the closest answer.

**0-19%   20-39%   40-59%   60-79%   80-99%   100-119%**





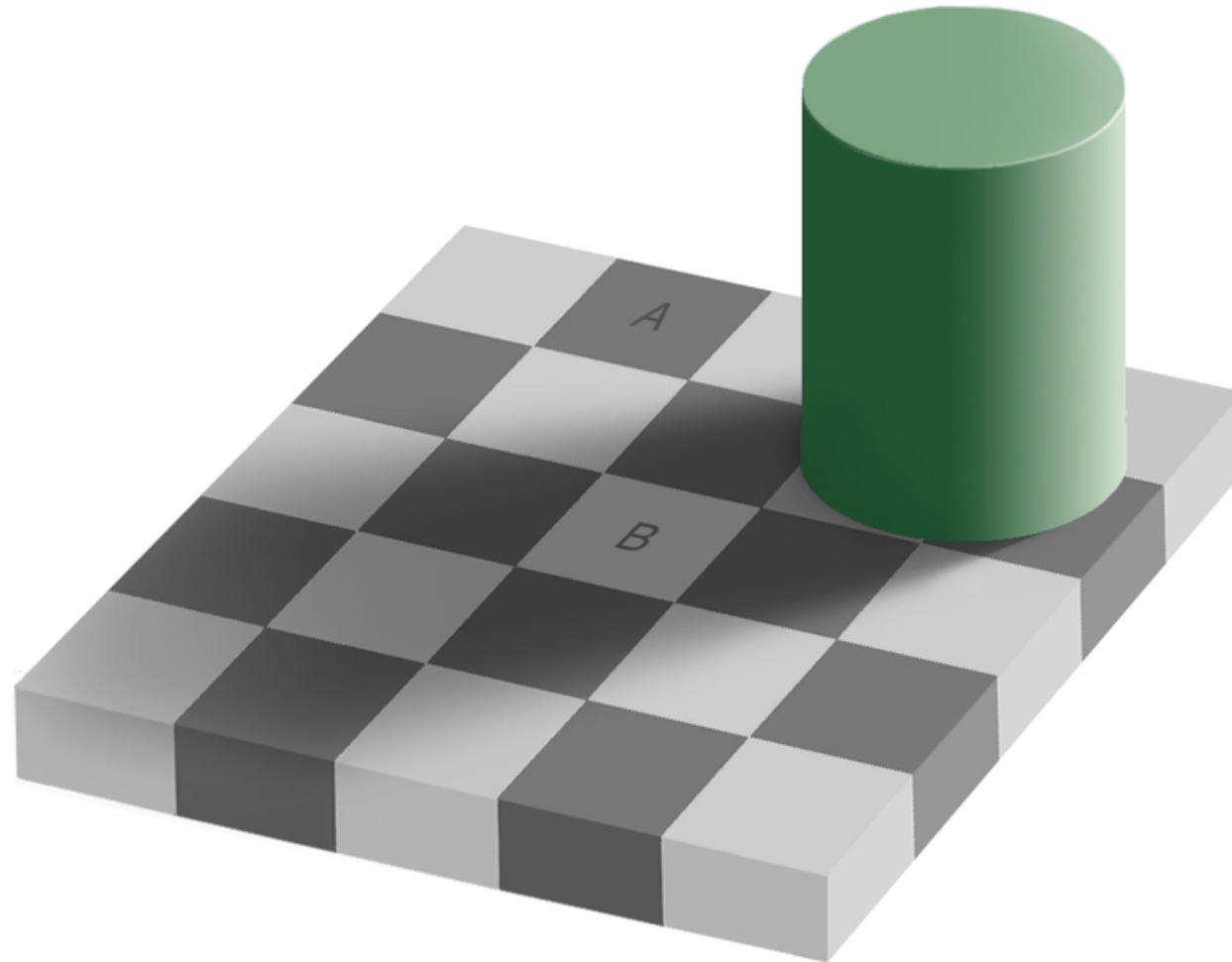


We have reason to suspect that Mark Smith is interested in terrorism. We have two reasons for thinking this.

- Smith is a member of the online white supremacist forum Lighting Strike (LS). The LS forum discusses topics such as an impending race war, race superiority, deep state conspiracies, and tips on using weapons, including improvised explosives and traps. From research done by experts in radicalisation it appears that about 20% of LS members think that terrorism is an acceptable means of achieving their political goals.
- In addition, Smith is also part of the “inner sanctum” of the LS forum, made up of the most active members of the forum. The inner sanctum members moderate the forum, decide who can join, and provide warnings about what can and cannot be said on the forum due to legal concerns. Members of the inner sanctum seem to be particularly concerned about perceived enemies such as deep state actors, immigrants, and the police. About 30% of members of the inner sanctum think terrorism is an acceptable means of achieving their political goals.

**Based on the information above, what is the probability that Smith thinks terrorism is an acceptable means of achieving his political goals?**

# Contrast Illusion



Edward H. Adelson

# Contrast Illusion (“proof”)



## Question

A bat and a ball together cost  
\$1.10.

The bat costs a dollar more than  
the ball.

How much does the ball cost?



Shane Frederick





## In COVID's shadow, global terrorism goes quiet. But we have seen this before, and should be wary

August 14, 2020 6.08am AEST

Alaa Al-Marjani/AAP

Email

Twitter 9

Facebook 133

LinkedIn

Print

Have we flattened the curve of global terrorism? In our COVID-19-obsessed news cycle stories about terrorism and terrorist attacks have largely disappeared. We now, though, understand a little more about how pandemics work.

And ironically, long before the current pandemic, the language of epidemiology proved helpful in understanding by analogy the way in which terrorism works as a phenomenon that depends on social contact and exchange, and expands rapidly in an opportunistic fashion when defences are lowered.

### Terrorism goes quiet - but we've seen this before

#### Author



**Greg Barton**

Chair in Global Islamic Politics,  
Alfred Deakin Institute for  
Citizenship and Globalisation,  
Deakin University

#### Disclosure statement

Greg Barton is engaged in a range of projects working to understand and

# What do you think of this argument?



In COVID's shadow, global terrorism has gone quiet but we should still be wary. Having lost its physical caliphate, Islamic State appears to have lost its capacity, if not its willingness, to launch attacks around the world well beyond conflict zones. But we have seen this happen before. The September 11 attacks in 2001 were followed by a wave of attacks around the world. Since 2005, except for the Charlie Hebdo shootings in Paris in January 2015, al-Qaeda has been prevented from launching any major attacks in western capitals. Then in 2013, Islamic State emerged and brought a new wave of attacks. While it's tempting to conclude that the ending of the current wave of international terrorist attacks by IS is due largely to the ending of the physical caliphate in Syria and Iraq, and a concomitant collapse of capacity, the reality is more complex.

The parallels with the epidemiology of viruses are striking. Terrorism works as a phenomenon that depends on social contact and exchange and expands rapidly in an opportunistic fashion when defences are lowered. Reasoning by analogy is imperfect, but it can be a powerful way of prompting reflection. The importance of this cannot be underestimated as intelligence failures in counterterrorism, like poor political responses to pandemics, are in large part failures of imagination. It is true we have successfully dealt with two waves of global terrorist attacks over the past two decades, but we have not dealt successfully the underlying source of infections. In fact, we have contributed, through military campaigns, to weakening the body politic of host countries in which groups like al-Qaeda, IS and other violent extremist groups have a parasitic presence. We now need to face the inconvenient truth that toxic identity politics and the tribal dynamics of hate have infected western democracies and that eliminating the viral spread of hateful extremism is extremely hard.

- How do you typically go about evaluating reasoning?
- What do you think makes reasoning good?
- What makes it poor?
- What are some of the things reasoning needs to successfully do?
- What are some of the things it should avoid doing?
- How do you communicate to colleagues your issues with their reasoning?

## **1.1 Why is evaluating reasoning difficult?**

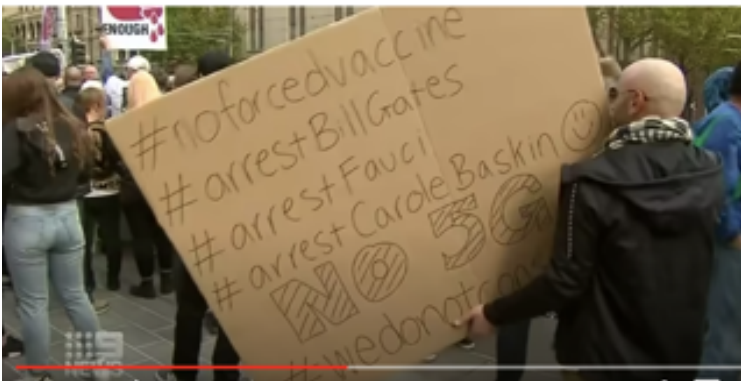
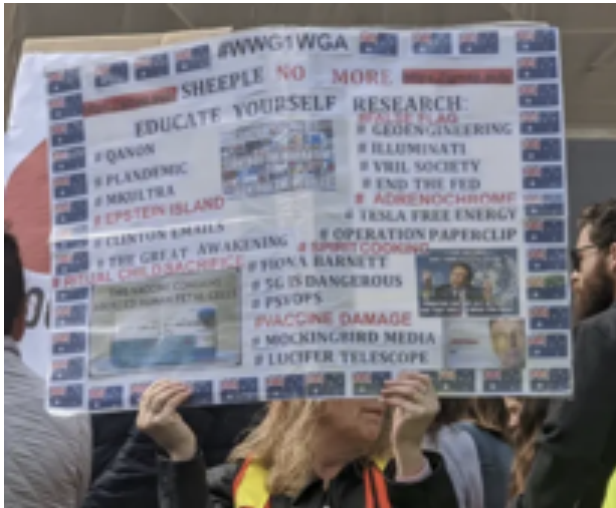




**Donald J. Trump**  @realDonaldTrump · May 26

**There is NO WAY (ZERO!) that Mail-In Ballots will be anything less than substantially fraudulent. Mail boxes will be robbed, ballots will be forged & even illegally printed out & fraudulently signed. The Governor of California is sending **Ballots** to millions of people, anyone.....**

 [Get the facts about mail-in ballots](#)





ARCHIVE

CURRENT ISSUE

THI

Adjust

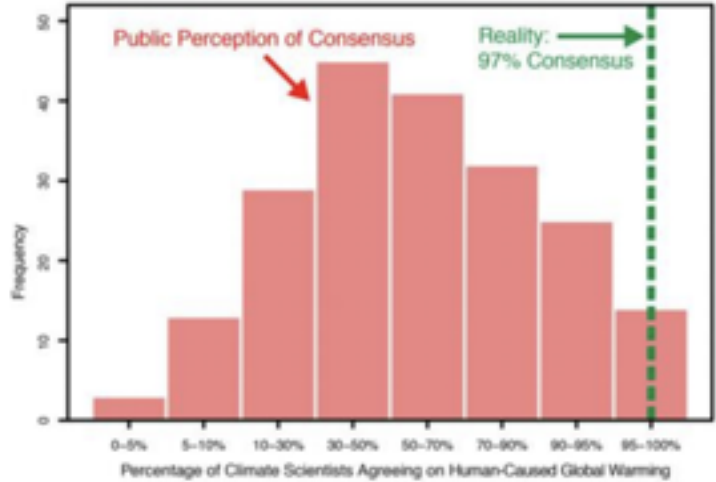
Share

### A Letter on Justice and Open Debate

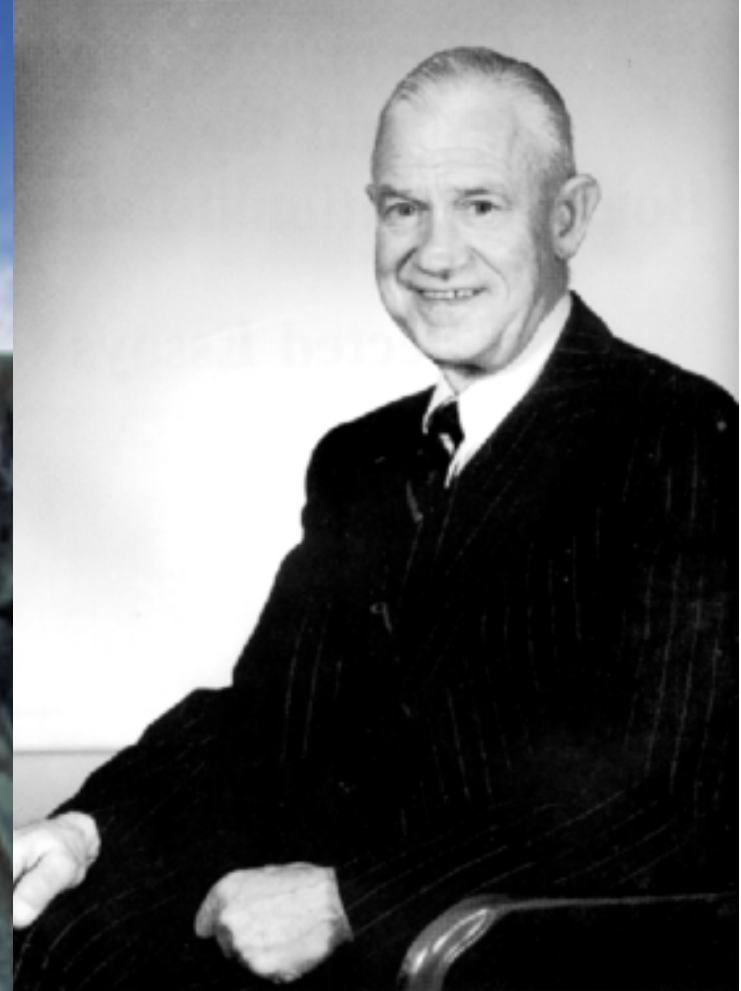
July 7, 2020

The below letter will be appearing in the Letters section of the magazine's October issue. We welcome responses at [letters@harpers.org](mailto:letters@harpers.org)

Our cultural institutions are facing a moment of trial. Powerful protests for racial and social justice are leading to overdue demands for police reform, along with wider calls for greater equality and inclusion across our society, not least in higher education, journalism, philanthropy, and the arts. But this needed reckoning has also intensified a new set of moral attitudes and political commitments that tend to weaken our norms of open debate and toleration of differences in favor of ideological conformity. As we applaud the first development, we also raise our voices against the second. The forces of illiberalism are



# A perennial problem





# Analytic Standards and Rigour

UNCLASSIFIED

Office of the Director of National Intelligence

Rating Scale  
for  
Evaluating Analytic Tradecraft Standards

## 6. USES CLEAR AND LOGICAL ARGUMENTATION.

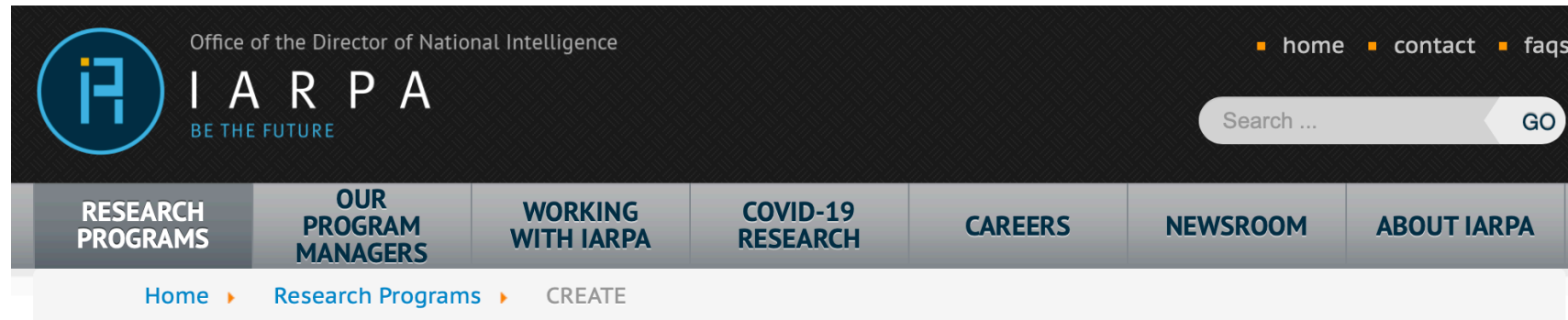
Definition: Analytic products should present a clear main analytic message up front. Products containing multiple judgments should have a main analytic message that is drawn collectively from those judgments. All analytic judgments should be effectively supported by relevant intelligence information and coherent reasoning.

Language and syntax should convey meaning unambiguously. Products should be internally consistent and acknowledge significant supporting and contrary information affecting judgments.

Poor	Fair	Good	Excellent
(1) Lacks a main analytic message; <b>OR</b> (2) Does not support analytic judgments with relevant evidence or undermines them by using flawed logic; <b>OR</b> (3) Often uses unclear language or uses a structure that is not easily understood.	(1) Presents a main analytic message; <b>BUT</b> (2) Does not combine evidence, context, and assumptions effectively to support analytic judgments or uses weak logic; <b>OR</b> (3) Sometimes uses unclear language or a structure that at times is not easily understood.	(1) Presents a prominent and clear main analytic message; <b>AND</b> (2) Presents clear reasoning with no flaws in logic and effectively combines evidence, context, and assumptions to support analytic judgments; <b>AND</b> (3) Uses clear language and a structure that displays a logical flow appropriate for the argument being presented.	Satisfies “good” criteria; <b>AND</b> (1) Addresses any inconsistent or contrary information in a way that reconciles it with analytic judgments; <b>OR</b> (2) Demonstrates notable skill or sophistication in combining evidence, context, and assumptions convincingly to support analytic judgments.

Attribute Description	Indicators of...		
	LOW Rigor	MODERATE Rigor	HIGH Rigor
<b>Hypothesis Exploration</b> The construction and evaluation of potential explanations for collected data.	<ul style="list-style-type: none"> <li>– Little or no consideration of alternatives to primary or initial hypotheses.</li> <li>– Interpretation of ambiguous or conflicting data such that they are compatible with existing beliefs.</li> <li>– Fixation or knowledge shielding behaviors.</li> </ul>	<ul style="list-style-type: none"> <li>– Some consideration of how data could support alternative hypotheses.</li> <li>– An unbalanced focus on a probable hypothesis or a lack of commitment to any particular hypothesis.</li> </ul>	<ul style="list-style-type: none"> <li>– Significant generation and consideration of alternative explanations via the direct evaluation of specific hypotheses.</li> <li>– Incorporation of “outside” perspectives in generating hypotheses.</li> <li>– Evolution and broadening of hypothesis set beyond an initial framing.</li> <li>– Ongoing revision of hypotheses as new data are collected.</li> </ul>
<b>Information Search</b> The focused collection of data bearing upon the analysis problem.	<ul style="list-style-type: none"> <li>– Failure to go beyond routine and readily available data sources.</li> <li>– Reliance on a single source type or on data that are far removed from original sources.</li> <li>– Dependence upon “pushed” information, rather than on actively collected information.</li> <li>– Use of stale or dated source data.</li> </ul>	<ul style="list-style-type: none"> <li>– Collection from multiple data types or reliance on proximal sources to support key findings.</li> <li>– Some active information seeking.</li> </ul>	<ul style="list-style-type: none"> <li>– Collection of data from multiple source types in addition to the use of proximal sources for all critical inferences.</li> <li>– Exhaustive and detailed exploration of data in the relevant sample space.</li> <li>– Active approach to information collection.</li> </ul>
<b>Information Validation</b> The critical evaluation of data with respect to the degree of agreement among sources.	<ul style="list-style-type: none"> <li>– General acceptance of information at face value, with little or no clear establishment of underlying veracity.</li> <li>– Lack of convergent evidence.</li> <li>– Poor tracking and citation of original sources of collected data.</li> </ul>	<ul style="list-style-type: none"> <li>– Use of heuristics to support judgements of source integrity, e.g. relying on sources that have previously proven to be consistently accurate.</li> <li>– A few “key” high-quality documents are relied on heavily.</li> <li>– Recognizes and highlights inconsistencies between sources.</li> </ul>	<ul style="list-style-type: none"> <li>– Systematic and explicit processes employed to verify information and to distinguish facts from judgements.</li> <li>– Seeks out multiple, independent sources of converging evidence.</li> <li>– Concerned both with consistency between sources and with validity and credibility within a given source.</li> </ul>
<b>Stance Analysis</b> The evaluation of collected data to identify the relative positions of sources with respect to the broader contextual setting.	<ul style="list-style-type: none"> <li>– Little consideration of the views and motivations of source data authors.</li> <li>– Recognition of only clearly biased sources or sources that reflect a well-defined position on an issue.</li> </ul>	<ul style="list-style-type: none"> <li>– Perspectives and motivations of authors are considered and assessed to some extent.</li> <li>– Incorporates basic strategies to compare perspectives of different sources, e.g. by dividing issues into “for” or “against” positions.</li> </ul>	<ul style="list-style-type: none"> <li>– Involves significant research into, or leverages a preexisting knowledge of, the backgrounds and views of key players and thought leaders.</li> <li>– May involve more formal assessments of data sources, e.g. via factions analysis, social network analysis, or deception analysis.</li> </ul>
<b>Sensitivity Analysis</b> The evaluation of the strength of an analytical assessment given possible variations in source reliability and uncertainty.	<ul style="list-style-type: none"> <li>– Explanations are appropriate and valid at a surface level.</li> <li>– Little consideration of critical “what if” questions, e.g., “What if a given data source turns out to be unreliable?” or “What if a key prediction does not transpire as anticipated?”</li> </ul>	<ul style="list-style-type: none"> <li>– Considers whether being wrong about some inferences would influence the overall best explanation for the data.</li> <li>– Identifies the boundaries of applicability for an analysis.</li> </ul>	<ul style="list-style-type: none"> <li>– Goes beyond simple identification to specify the strength of explanations and assessments in the event that individual supporting evidence or hypotheses were to prove invalid or unreliable.</li> <li>– Specifies limitations of the analysis, noting the most vulnerable explanations or predictions on which the analysis is at risk of erring.</li> </ul>
<b>Information Synthesis</b> The extent to which an analyst goes beyond simply collecting and listing data in “putting things together” into a cohesive assessment.	<ul style="list-style-type: none"> <li>– Little insight with regard to how the analysis relates to the broader analytical context or to more long-term concerns.</li> <li>– Lack of selectivity, with the inclusion of data or figures that are disconnected from the key arguments or central issues.</li> <li>– Extensive use of lists or the restatement of material copied directly from other sources with little reinterpretation.</li> </ul>	<ul style="list-style-type: none"> <li>– Explicit, though perhaps not systematic, efforts to develop the analysis within a broader framework of understanding.</li> <li>– Depiction of events in context and framing of key issues in terms of tradeoff dimensions and interactions.</li> <li>– Provides insight beyond what is available in the collected data.</li> </ul>	<ul style="list-style-type: none"> <li>– Extracted and integrated information in terms of relationships rather than components and with a thorough consideration of diverse interpretations of relevant data.</li> <li>– Re-conceptualization of the original task, employing cross-checks on abstractions.</li> <li>– Performed by individuals who are “reflexive” in that they are attentive to the ways in which their cognitive processes may have hindered effective synthesis.</li> </ul>
<b>Specialist Collaboration</b> The extent to which substantive expertise is integrated into an analysis.	<ul style="list-style-type: none"> <li>– Minimal direct collaboration with experts.</li> <li>– Little if any on-topic, “outside” expertise is accessed or sought out directly.</li> </ul>	<ul style="list-style-type: none"> <li>– Involves some direct interaction with experts, though usually via readily available specialists.</li> <li>– Expertise is drawn from within preexisting personal or organizational networks.</li> </ul>	<ul style="list-style-type: none"> <li>– Independent experts in key content areas are identified and consulted.</li> <li>– Efforts to go beyond a “core network” of contacts to seek out domain-relevant experts, with additional resources and “political capital” potentially expended to gain access to such specialist expertise.</li> </ul>
<b>Explanation Critiquing</b> The critical evaluation of the analytical reasoning process as a whole, rather than in the specific details.	<ul style="list-style-type: none"> <li>– Few if any instances of alternative or “outside” criticisms being considered.</li> <li>– Reliance on preexisting channels of critiquing, primarily those supervisory.</li> </ul>	<ul style="list-style-type: none"> <li>– Brings alternative perspectives to bear in critiquing the overall analytical process.</li> <li>– Leverages personal or organizational contacts to examine analytical reasoning, e.g. by way of peer analysts, proxy decision makers, etc.</li> </ul>	<ul style="list-style-type: none"> <li>– Familiar as well as independent perspectives have examined the chain of analytical reasoning, explicitly identifying which inferences are stronger and weaker.</li> <li>– Use of formal methods such as “red teams” or “devil’s advocacy” to challenge and vet hypotheses and explanations.</li> <li>– Expenditure of capital, political or otherwise, in critiquing the analytical process.</li> </ul>





## Crowdsourcing Evidence, Argumentation, Thinking and Evaluation (CREATE)

### CREATE BETTER REASONING

The CREATE program seeks to develop, and experimentally test, systems that use crowdsourcing and structured analytic techniques to improve analytic reasoning. These systems will help people better understand the evidence and assumptions that support—or conflict with—conclusions. Secondly, they will also help users better communicate their reasoning and conclusions.

The CREATE program is envisioned as a 4.5-year effort that is intended to begin in September 2016. Phase 1 of the program lasted 23 months, Phase 2 will last 17 months and Phase 3 will last 14 months.

**Join the CREATE study**

#### Program Manager

[Steve Rieber](#)

#### Program Information

[IARPA-BAA-15-11](#)

#### COVID-19 Related Research



#### Research Area(s)

- Social and behavioral sciences
- Informal reasoning
- Computer science
- Structured analytic techniques
- Crowdsourcing

## 1.2 RST method and training

## RST Steps

1. Reasoning review
2. Flaw identification
3. Overall Evaluation
4. Fixing flaws



- Logical fallacies, faulty assumptions, probabilistic misconceptions
  - » e.g., Testing rates and Covid-19 cases correlate, so testing causes Covid.
- Missing important citations
- Contradictions in assessments/products
  - » e.g., Warrant application contradicts other reporting.
- Basing judgements on evidence without considering alternatives
  - » e.g., Rejecting a sources as vexatious just because they appear to have a bias
- Introducing evidence that points different ways and not explaining why one side should be favoured over the other.

- Improve **review of reasoning** in analytic products
- Improve ability to diagnose and **fix reasoning flaws**
- Better understand **reasoning structure** and important types of arguments
- Created a **shared language** for effectively discussing reasoning
- Provide more useful and **insightful feedback**
- Expand your body of collective knowledge to help **champion analytic rigour** and reasoning quality



- 1. Fundamentals of good reasoning**
- 2. Reasoning structure**
- 3. Reasoning stratagems (for establishing claims)**
- 4. Identify reasoning flaws**
- 5. Assessing the impact of reasoning flaws**
- 6. Fixing flaws**
- 7. Providing feedback**
- 8. Consolidation**



Deliberate practice

» Regular quizzes

» Observable techniques

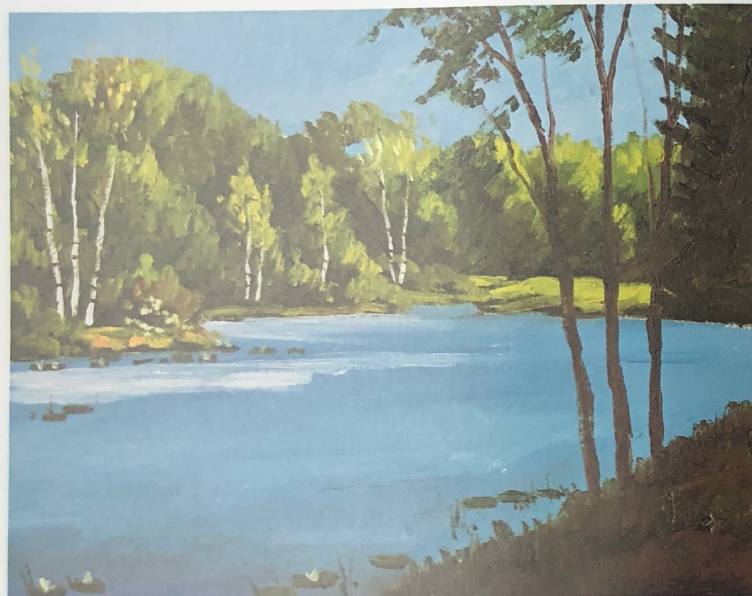


e.g., The writing technique of removing prepositions:

**"The project is likely to result in a minor population increase in the city from families relocating to the site from outside the community."**

**"The project will probably attract new families to the city."**

### 30. Observe More Colors than Blue in Water



#### PROBLEM

One of the greatest examples of how the subconscious mind dominates our analytical perception is how we see water. If you ask the average person what color water is, he will invariably say blue, and this is the way most amateurs paint it. For the same reason, skies are often painted an uninteresting, flat blue. Actually, water has little if any color. We think of it as blue because it often reflects a blue sky. Water acts as a mirror and the sharpness of the reflections on it is governed by the degree of agitation. Wind ripples—tiny waves caused in the water by the wind—are lighter than the surrounding reflections because they reflect light from the sky above. In this demonstration we have a classic example of the mind telling us that the answer to painting water is blue. If a person has not had

enough training to enable him to see values and colors in very undefined, subtle patterns, he invariably resorts to this cliché.

Notice also that the distance is too similar in color and value to the foreground trees that are superimposed over it. These same foreground trees are poorly designed, too evenly spaced. The pine tree in the upper right has a repetitious saw-tooth edge. In painting birch trees, one has to be careful of the corny overstatement that causes some representational painting to be regarded with disfavor. The attempt in the foreground to say pond-lilies is contrived in placement and handling, and the right foreground—which should not come exactly to the center—has not been given enough sensitive detail.



SUMMER REFLECTIONS. Oil on canvas, 24" x 30".

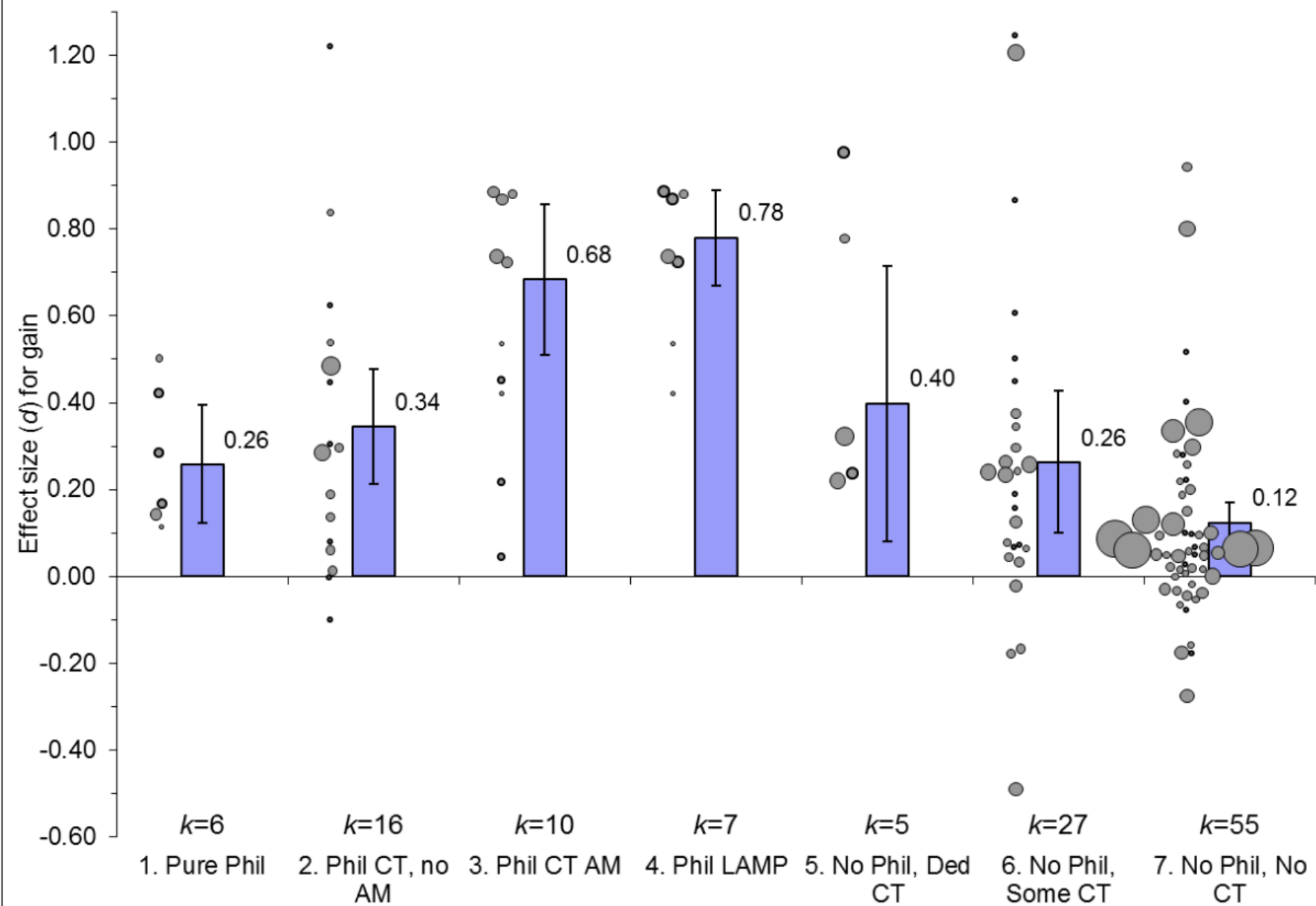
#### SOLUTION

There are absolutely beautiful patterns and colors in water if you know what to look for and how to see it. The weeds, algae, scum, and plants offer wonderful chances for color supplements such as in the yellow growth on the left and purple pickerelweed in the far right. I usually paint water slightly agitated so the mirror reflection is not as sharp as the image above it. Wind ripples can have color, if we have color in the sky above to justify it. Here we have a cold pink with a light cerulean blue painted into it to give us a neutral unobtainable in any other way.

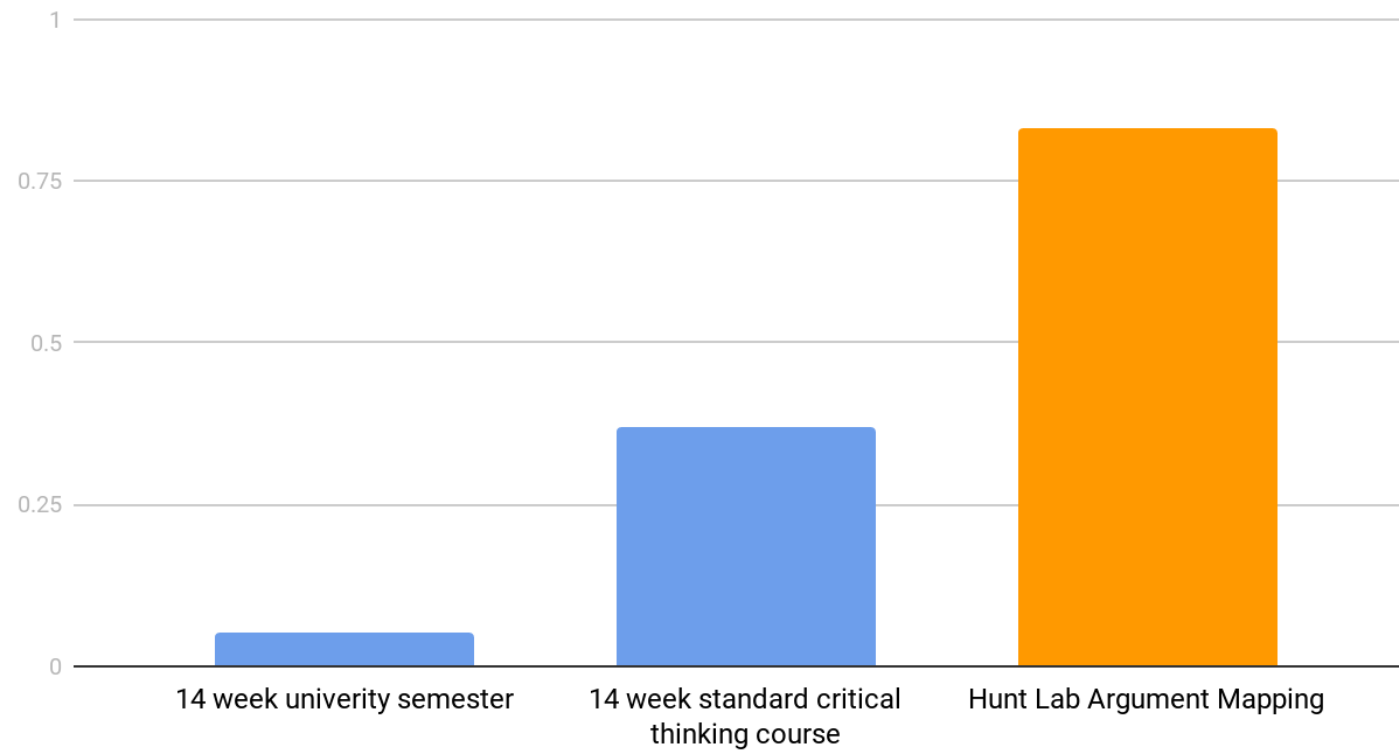
The early morning haze gave me a chance to use more atmospheric colors in the distance, as well as an opportunity to actually show the rays of sunlight. I have played

light foliage in the distant plane against the pine tree on the right. Note also the varied edges of the pine silhouette. Every opportunity should be used to introduce warm colors into a green painting. I have dealt with this more in Key No. 28, but I do want to call your attention to this additional example in the island on the left. The birches mentioned in the problem are here also, but handled softly so that they fit in unobtrusively. Notice also how the ones on the left side are made just a bit higher than the mass of trees behind them. The greater detail in the right foreground, in contrast to the only suggested detail in the distance, is always recommended to help the illusion of depth.





## Critical thinking gains



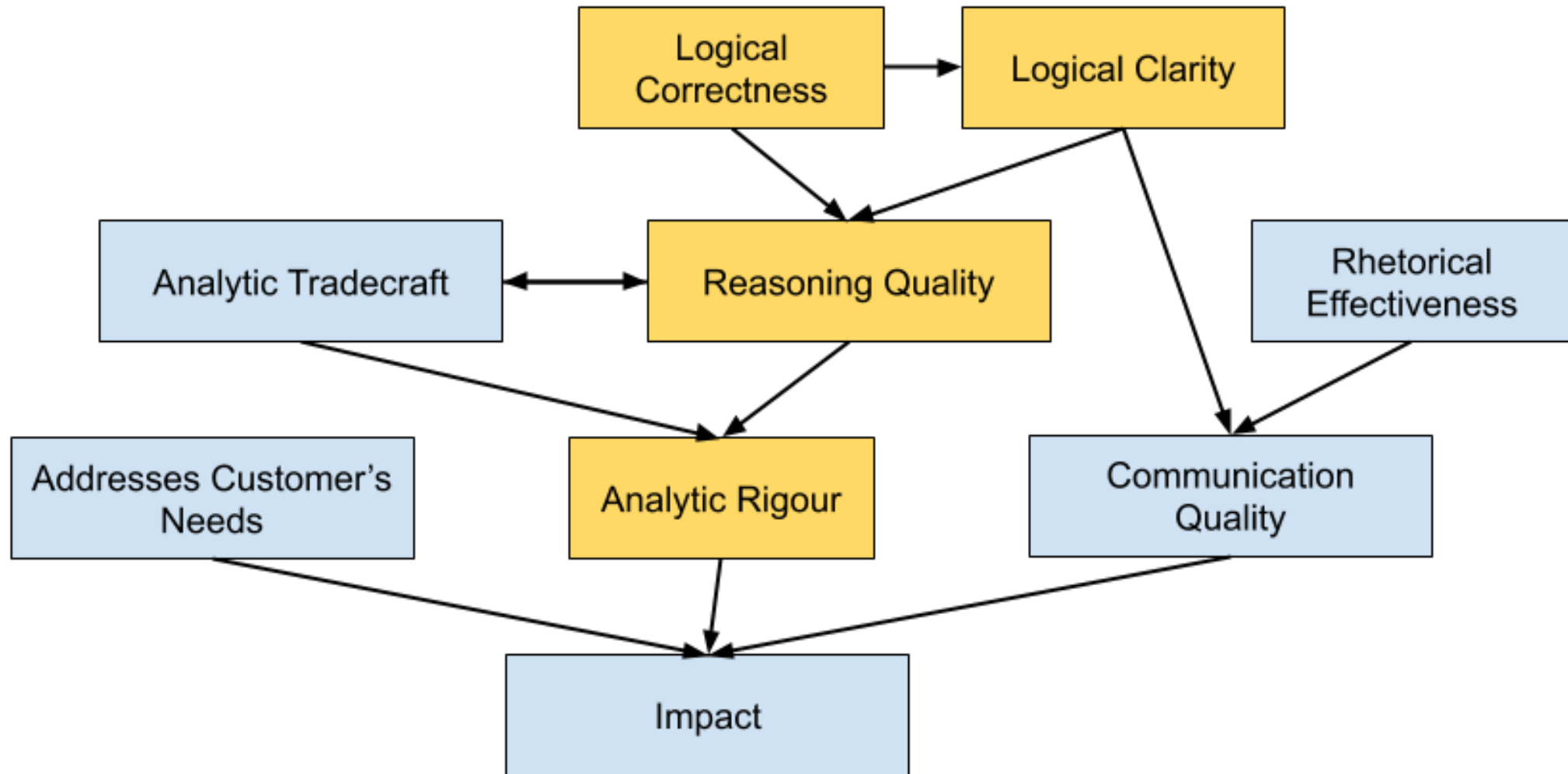
## 1.3 Reasoning and Rigour

# What is analytic rigour?

- Hypothesis exploration
- Information search
- Information validation
- Source evaluation
- Assessment of uncertainty
- Information synthesis
- Specialist collaboration
- Explanation critiquing
- Accountability

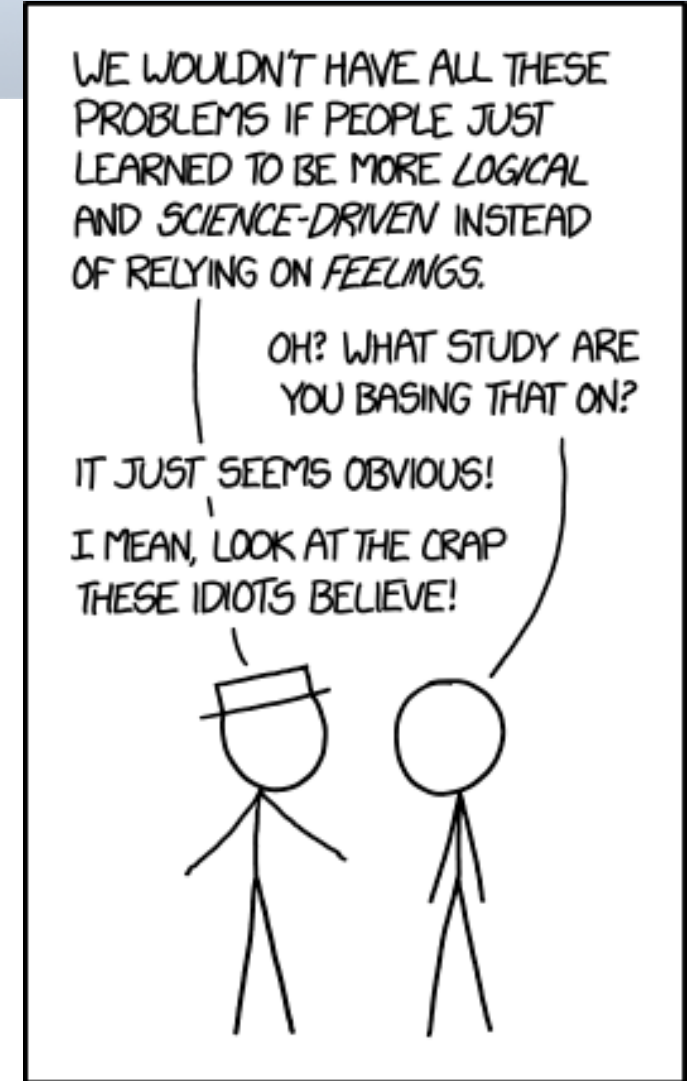






# Reasoning Skills

- Assessing the strength of reasons
- Analysing the structure of arguments
- Describing claims and hypotheses precisely
- Identifying assumptions
- Generating alternatives, counterfactuals, and counterexamples
- Making important distinctions and noticing similarities
- Critical self-reflection and avoiding biases
- Identifying reasoning flaws



Source: <https://xkcd.com/1901/>

Closely tied to the accountability/transparency of your assessments and efficiency of communication.

If your reasoning is clear and explicit:

- Reviewers of your assessments will have less basic/clarificatory/"nagging" questions and the discussion will be more about the core issues
- It will make your team's life easier. Someone can pick up your work and understand the justification
- You will be well prepared for internal and external checks of rigour and accountability

- "Good", correct reasoning doesn't guarantee insight
  - You can easily reason correctly about trivial things
- Some argue that too much emphasis on rigour will reduce insight
  - This is only a problem if you try to "game" rigour, and our approach should help prevent that
- When reasoning about complicated matters, the more rigorous your reasoning is, the more insights you will be able to achieve.
  - e.g., Lack of rigour encourages us to accept our first ideas, good or not

**Questions?**