

From Psychic Claims to Science: Testing Psychic Phenomena with Statistics



Jessica Utts

Department of Statistics

University of California, Davis

<http://anson.ucdavis.edu/~utts>

jmutts@ucdavis.edu



TWO QUESTIONS

- Question 1:

What do you really know to be true and how do you know it? “Private Knowing”

- Question 2:

What do we all agree is true and why do we agree that it is true? “Public Knowledge”



Private Knowing – Some Responses

- Experience

"I know I love my kids."

- "Gut Feeling"

"I know psychic phenomena are real."

- Belief in "experts"

"I know the speed of light is 186,000 miles per second."



Public Knowledge – Some Responses

- Shared Experience and Beliefs

"On a hot day it's cooler in the shade than in the sun."

- Physical and Biological Laws and Observations

"The earth rotates around the sun."

- Statistical Studies

"Aspirin helps prevent heart attacks."

"Nicotine patches help people quit smoking."



QUESTION

HOW DO WE MOVE:

- FROM PRIVATE KNOWING TO PUBLIC KNOWLEDGE?
- FROM ANECDOTES TO SCIENTIFIC CONCLUSIONS?



ANSWER

Moving from anecdotes to knowledge requires:

- Creating testable hypotheses
- Designing suitable experiments
- Analyzing results
- Creating new hypotheses

And eventually...

- Accumulating sufficient [statistical] evidence, and an explanation.



Psi/Psychic/ESP/Anomalous Cognition

Having information that could not have been gained through the known senses.

- Telepathy: Info from another person
- Clairvoyance: Info from another place
- Precognition: Info from the future

For proof -> Source isn't important.

For explanation -> Source is important.



Controlled experiments to Test ESP

Crucial elements:

1. Safeguards to rule out cheating or ordinary means of communication
2. Knowledge of the probabilities of various outcomes by chance alone

Examples... are these okay?

1. I am thinking of a number from 1 to 5. Guess it.
2. My assistant down the hall has shuffled a deck of cards and picked one. What suit is it?

Free response ESP experiments meeting the conditions above:

- One type (originally done by US Government) is called *Remote Viewing*
- Similar type of experiment is called *ganzfeld*.
- I don't have time to explain both; will explain remote viewing and show results for both types.



Controlled experiments to Test ESP

Crucial elements:

1. Safeguards to rule out cheating or ordinary means of communication
2. Knowledge of the probabilities of various outcomes by chance alone

Examples... are these okay?

1. I am thinking of a number from 1 to 5. Guess it.

Not OK. Doesn't rule out subtle cues as ordinary means of communication, and probabilities are not known – certain numbers are more likely to be the ones thought of as well as guessed.

2. My assistant down the hall has shuffled a deck of cards and picked one. What suit is it?

OK. Fits both criteria (assuming cards are well-shuffled).

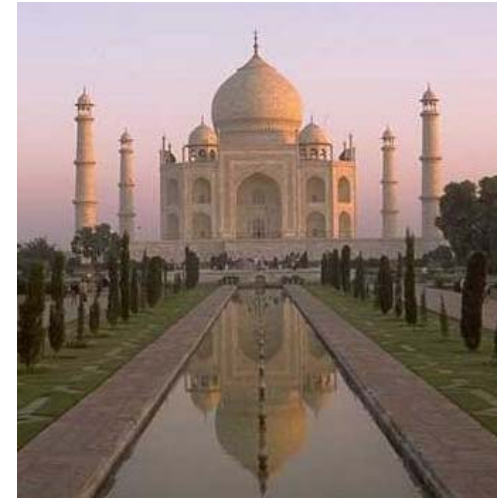
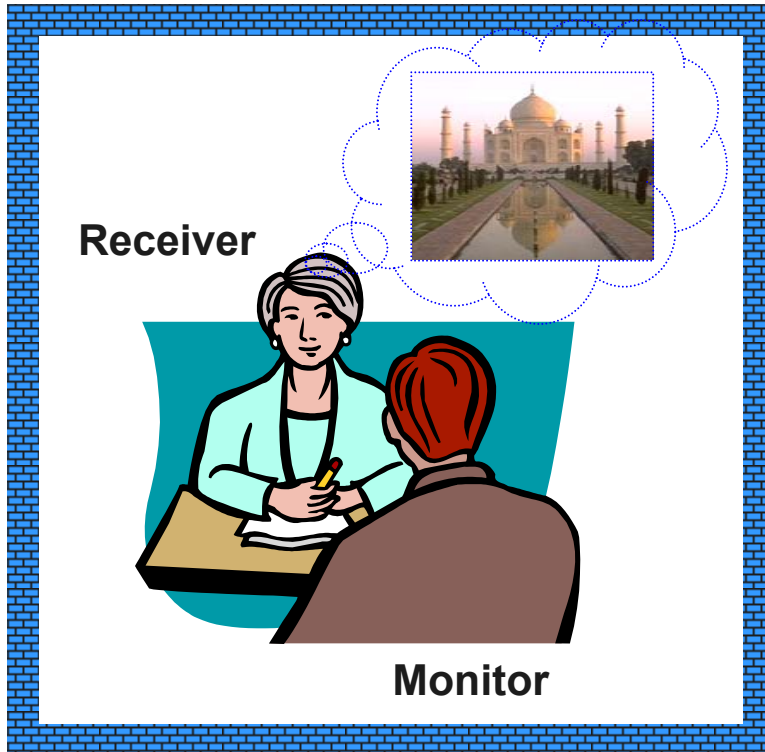
Free response ESP experiments meeting the conditions above:

- One type (originally done by US Government) is called *Remote Viewing*
- Similar type of experiment is called *ganzfeld*.
- I don't have time to explain both; will explain remote viewing and show results for both types.

Remote Viewing Protocol

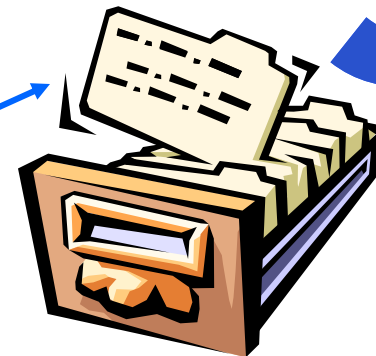
Special thanks to Dr. Edwin May for the next 4 slides

10:00



15 Minutes

10:05



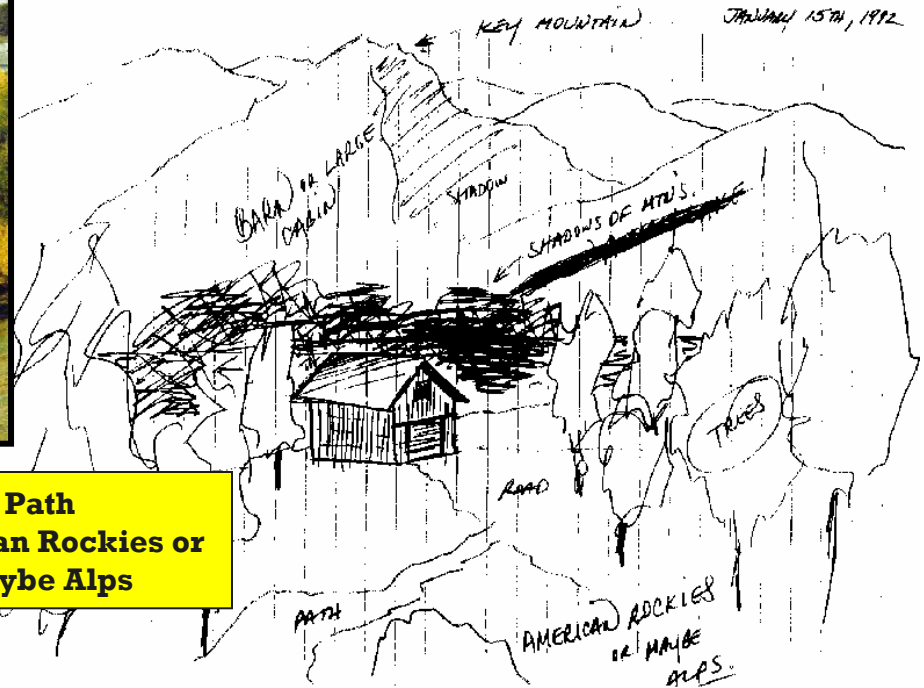
Example of an Excellent Match

(Experiment at SAIC/Stanford)

Target on left, Remote viewer's work on right



Key Mountain
Barn or Large Cabin
Shadow
Shadows of Mtns.
Trees
Road

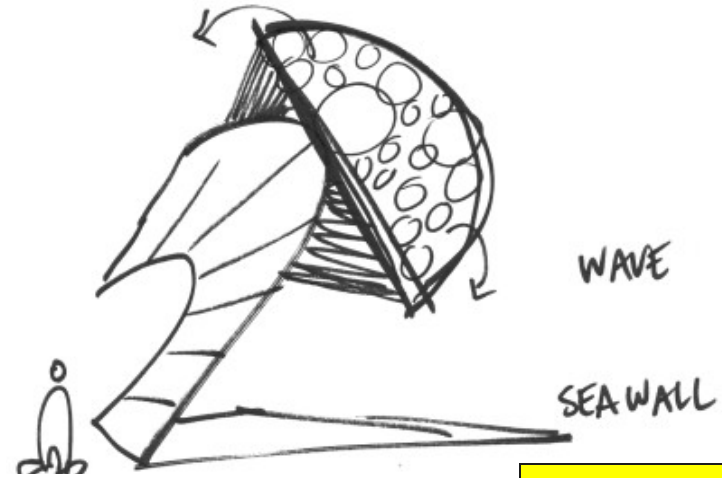
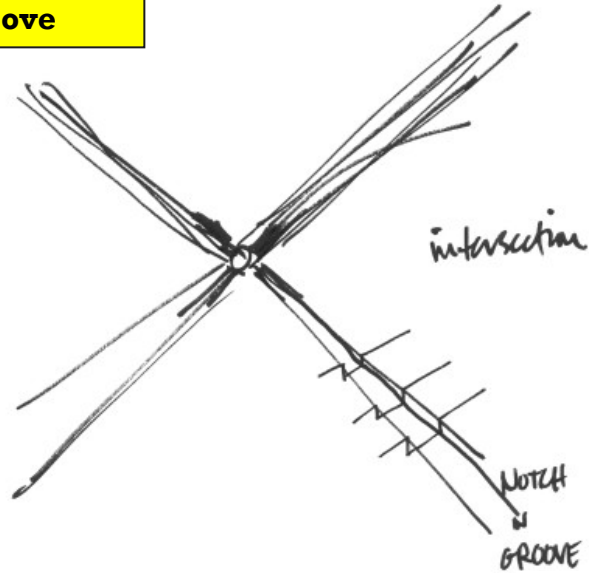


(Yellow boxes decipher what the remote viewer wrote!)

Path
American Rockies or
Maybe Alps

Typical Response – Novice

intersection,
notch, groove



LOOKING DOWN:

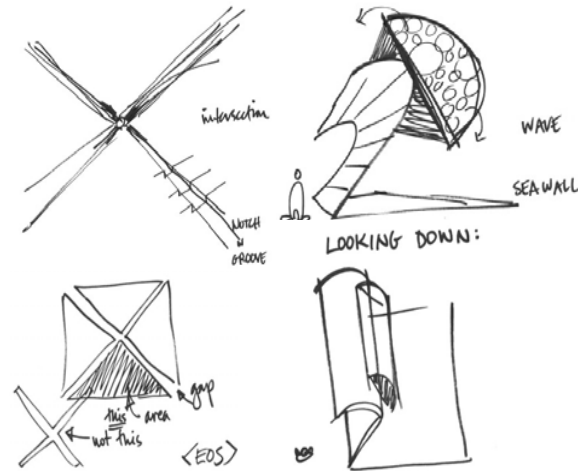
wave, sea wall

gap



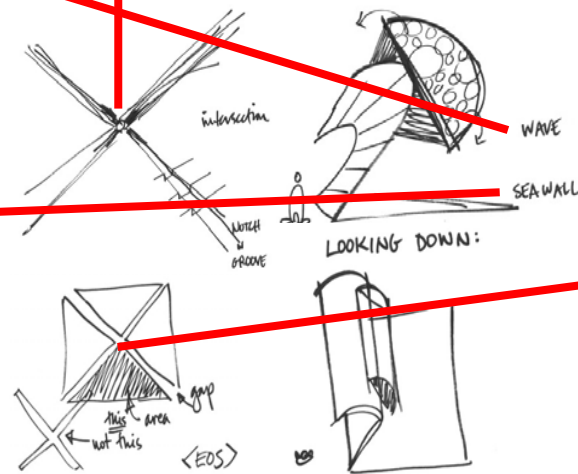
Rank-Order Analysis

Five choices – which matches best, 2nd, etc?



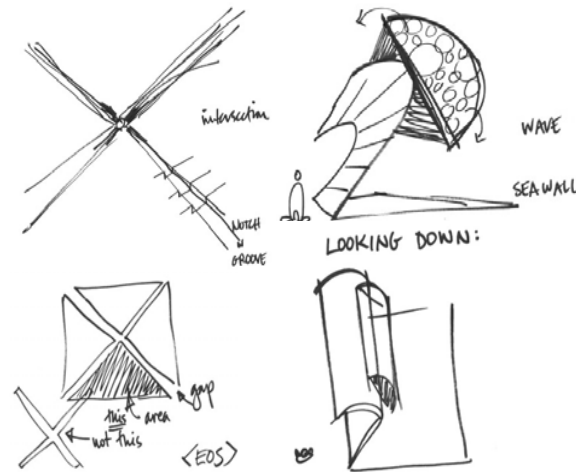
Which is the Actual Target?

Most have features that match the drawing!



Judge's rankings; Windmill is correct!

This was a 1st place match.





Analysis Methods

- Before the experiment, targets are arranged into packs of 4 or 5 (depending on lab) dissimilar choices.
- Before session begins a pack is randomly selected, then target within it.
- After the experiment, judge matches the response to the 4 (or 5) choices. Judge is blind to correct answer.
- Use either the *rank* assigned to correct target, *or* “direct hit” if it gets 1st place rank.
- Use sum or ranks (some labs), or number of direct hits (others), for entire experiment.

What Constitutes Evidence for Statistical Studies?



- Small *p-values* (probability of results as good as those seen or better, if just guessing)
- *Confidence intervals* showing similar effects in a variety of similar situations, labs, etc.
- Independent and conceptual replication
 - Remote viewing and ganzfeld, for instance



What is a confidence interval?

- There is an unknown, unmeasurable “true” value, such as proportion of smokers who would quit with nicotine patch or proportion of times correct target gets chosen by judge in ESP test
- Collect data using a *sample*. Using *sample proportion*, get interval that almost certainly covers *true proportion*.
- Works 95% of the time, so the interval is called a 95% confidence interval.



Nicotine Patch Example (*Seeing Through Statistics*)

- 240 smokers, 120 randomly assigned to wear nicotine patch, 120 to placebo.
- After 8 weeks, 55 nicotine patch users had quit; $55/120 = .46$ or 46%. 95% *C.I. for population is 37% to 55%.*
- Only 20% of placebo patch users quit; 95% *CI is 11% to 29%.*
- “Significant difference” p -value $< .001$.

Results of Free Response Experiments

(Used in 1995 report I wrote for U.S. Government)

Hit rates assume there were *four* choices; chance = 25%

U.S. Government Studies in Remote Viewing:

- SRI International (1970's and 1980's)
966 trials, p-value = 4.3×10^{-11} , **hit rate = 34%**, 2-sided 95% C.I. 31% to 37%
- SAIC
455 trials, p-value = 5.7×10^{-7} , **hit rate = 35%**, C.I. 30% to 40%

Ganzfeld:

- Psychophysical Research Laboratories, Princeton (1980's)
355 trials, p-value = .00005, **hit rate = 34.4%**, C.I. 29.4% to 39.6%
- University of Amsterdam, Netherlands (1990's)
124 trials, p-value = .0019, **hit rate = 37%**, C.I. 29% to 46%
- University of Edinburgh, Scotland (1990's)
97 trials, p-value = .0476, **hit rate = 33%**, C.I. 25% TO 44%
- Rhine Research Institute, North Carolina (1990's)
100 trials, p-value = .0446, **hit rate = 33%**, C.I. 24% to 42%

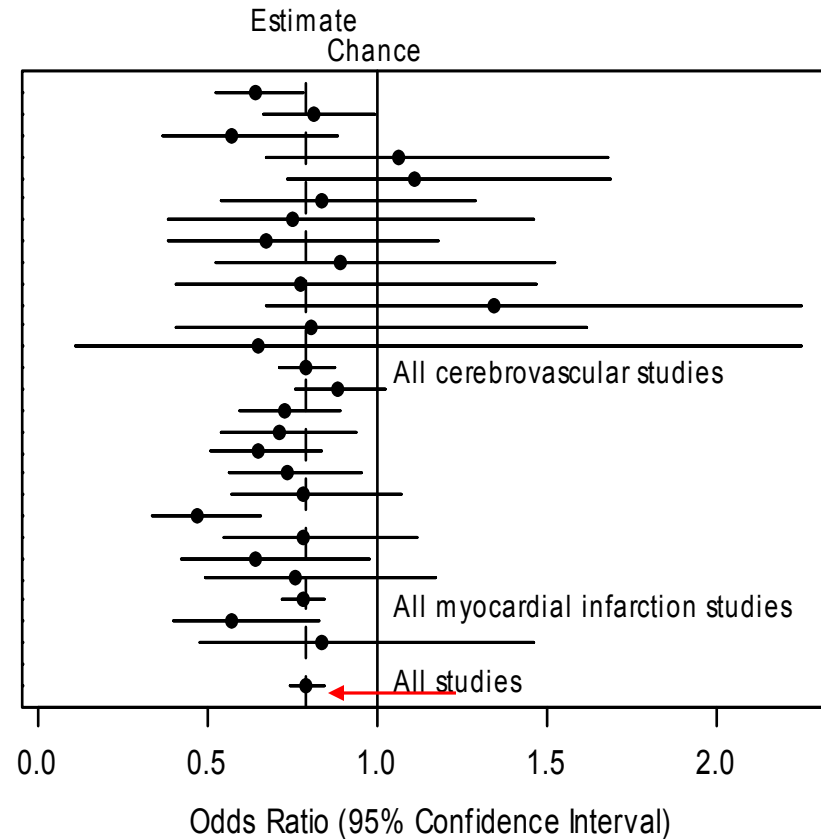


Non-controversial example: Aspirin and Recurring Vascular Disease

- Meta-analysis of 25 clinical trials on recurrence of heart attack or stroke when taking aspirin versus placebo.
- Outcome of interest: Odds ratio
- Odds of recurrence aspirin/placebo
- Chance \rightarrow Odds ratio = 1
 - 25 Studies, 5 with *p-value* < .01
 - Combined odds ratio of 0.75, represents 25% drop in recurrence rate of heart attacks

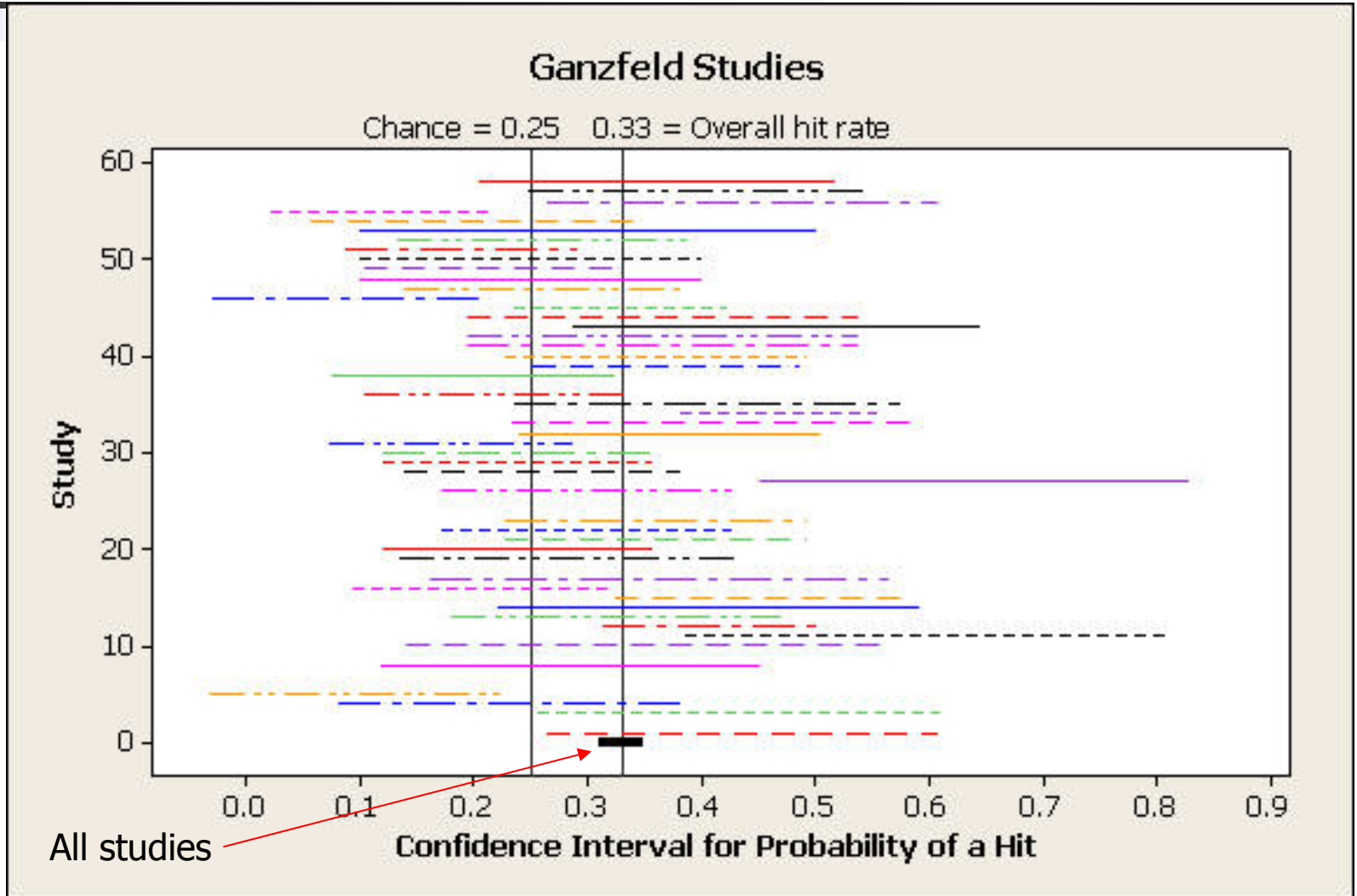
Confidence Intervals for Odds Ratio

Each line represents one study. Vertical lines at .75 (average for all studies) and 1 (value indicating no effect, just chance)



Ganzfeld Studies

58 Studies, overall hit rate = 33%





Quotes about aspirin studies

- The trials were very heterogeneous, including a range of ages, a range of different diseases, a range of treatments, and so on.
- Though such risk reductions might be of some practical relevance, however, they are surprisingly easy to miss, even in some of the largest currently available clinical trials. If, for example, such an effect exists, then even if 2000 patients were randomised there would be an even chance of getting a false negative results...that is, of failing to achieve convincing levels of statistical significance ($p < .01$).
- The main results were obtained from the principal investigators in most cases. In some trials the data obtained differed slightly from the data originally published.
- The final meeting of collaborators was supported not only by the [UK] Medical Research Council and Imperial Cancer Research Fund but also by the Aspirin Foundation, Rhone-Poulenc Sante, Reckitt and Colman, Bayer, Eli Lilly, Beechams, and the United Kingdom Chest, Heart and Stroke Association.

And... what was to prevent having pill analysed by local chemist?



Compare to Aspirin/Heart Attack Studies

How are anomalous cognition (ac) - remote viewing and ganzfeld - results different from aspirin results?

- If *same* standard applied, ac results are *much stronger*.
- The aspirin studies had *more opportunity* for fraud and experimenter effects than did the ac studies.
- The aspirin studies were *at least* as frequently funded and conducted by those with a *vested interest* in the outcome.
- Both used *heterogeneous* methods and participants.



Which makes me wonder...

- Why are millions of heart attack and stroke patients taking daily aspirin, but many people don't even know about the remote viewing and ganzfeld results?
- Why do many people who do know about them refuse to accept the evidence?



Establishing Public Knowledge

- How do we really know what we know?
- What roles do personal experience versus “objective” information play in what we think we know?
- Would *you* be more convinced by hundreds more statistical studies, or by one overwhelming personal experience?