You already know what this talk is about

Daniel Matheson

ACTSC 991 - Spring 2019

July 15th 2019

1. Introduction

2. Experiment Description

3. Discussion/Criticism

- We are refereeing Bem, D. J. (2011). Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of personality and social psychology* 100(3), 407.
- This is a paper about *Extra Sensory Perception (ESP)* in which Bem describes 9 different experiments he performed at Cornell to test for ESP.

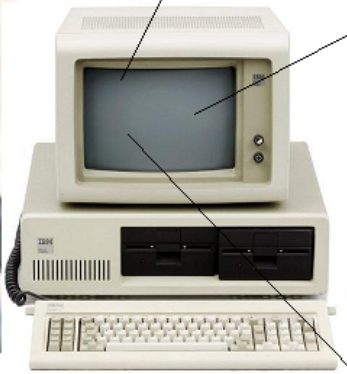
What is ESP (or psi)?

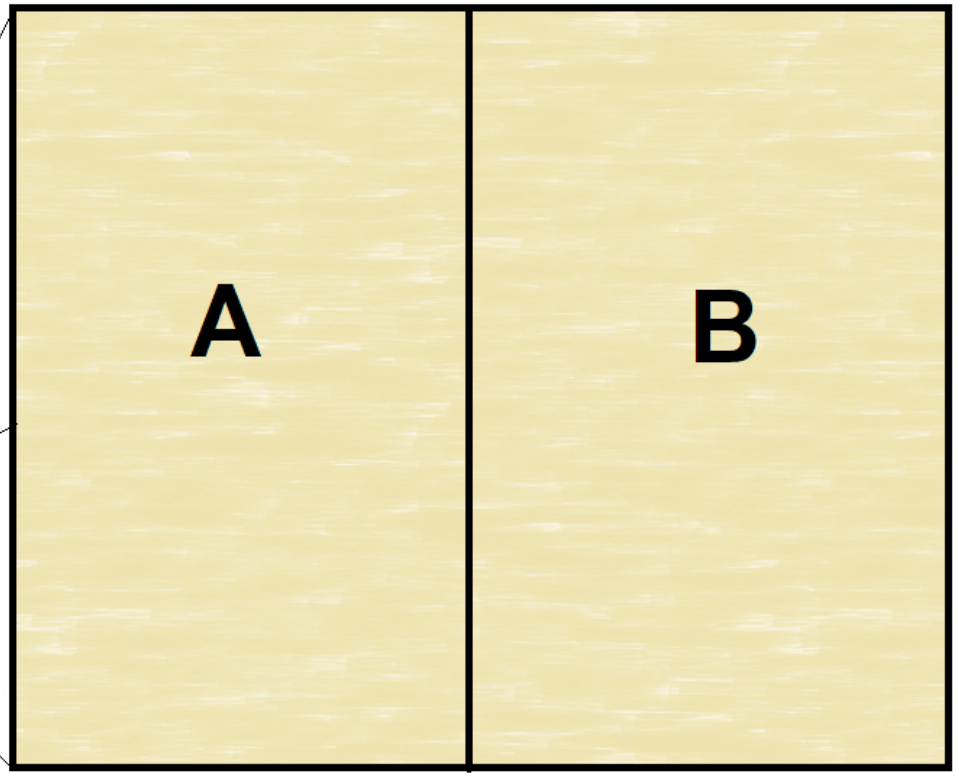
- From Bem: "The term **psi** denotes anomalous processes of information or energy transfer that are currently unexplained in terms of known physical or biological mechanisms"
- Two variants of psi of interest to Bem are:
  - Precognition: *consciously knowing* something is going to happen before it happens.
  - Premonition: *feeling* something is going to happen before it happens.

Experiment 1: Precognitive Detection of Erotic Stimuli

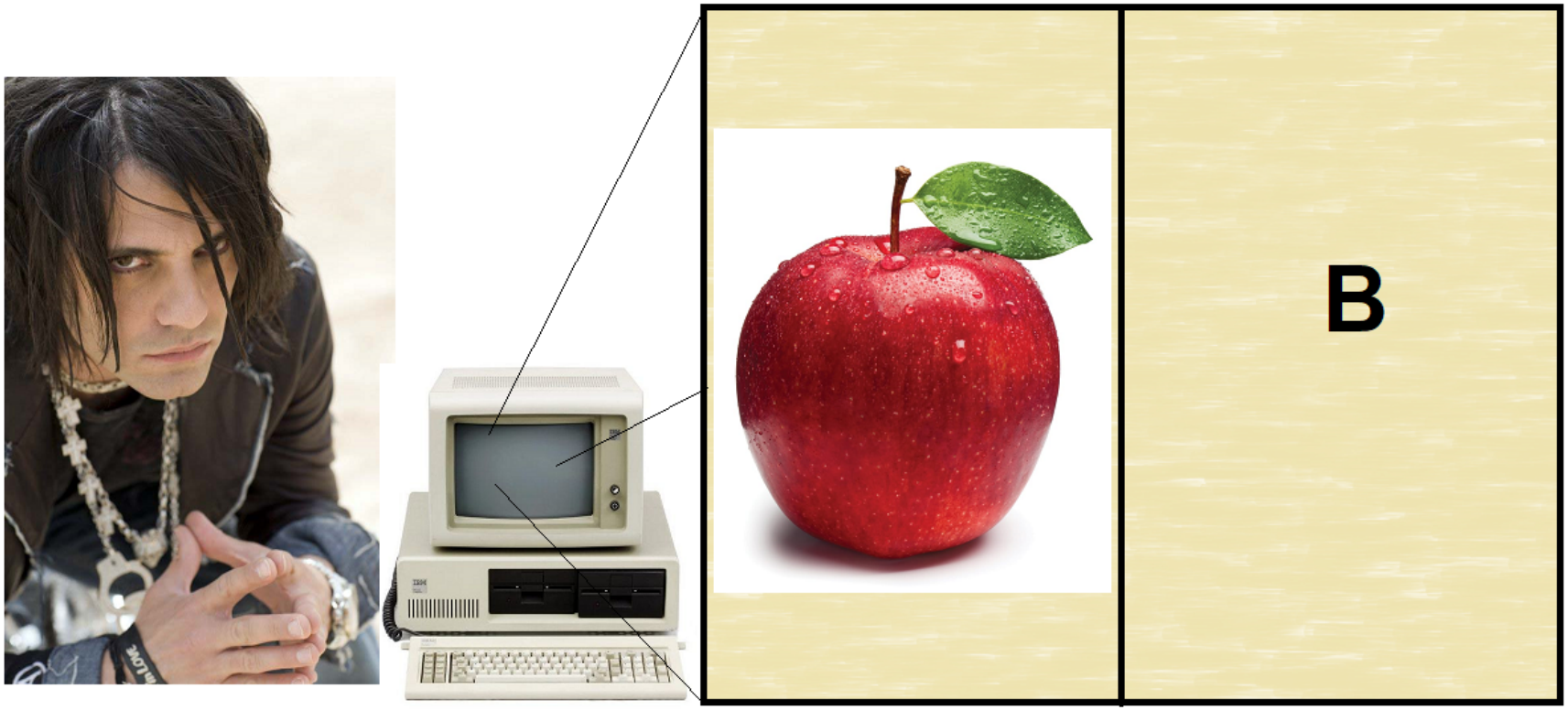
- In Bem's first experiment, Bem is attempting to prove there is evidence that *"our physiology can anticipate unpredictable erotic or negative stimuli before they occur"* $\Rightarrow$ **research question**.
- We will examine whether or not Bem's statistical methods actually answer this question adequately.

Experimental set-up

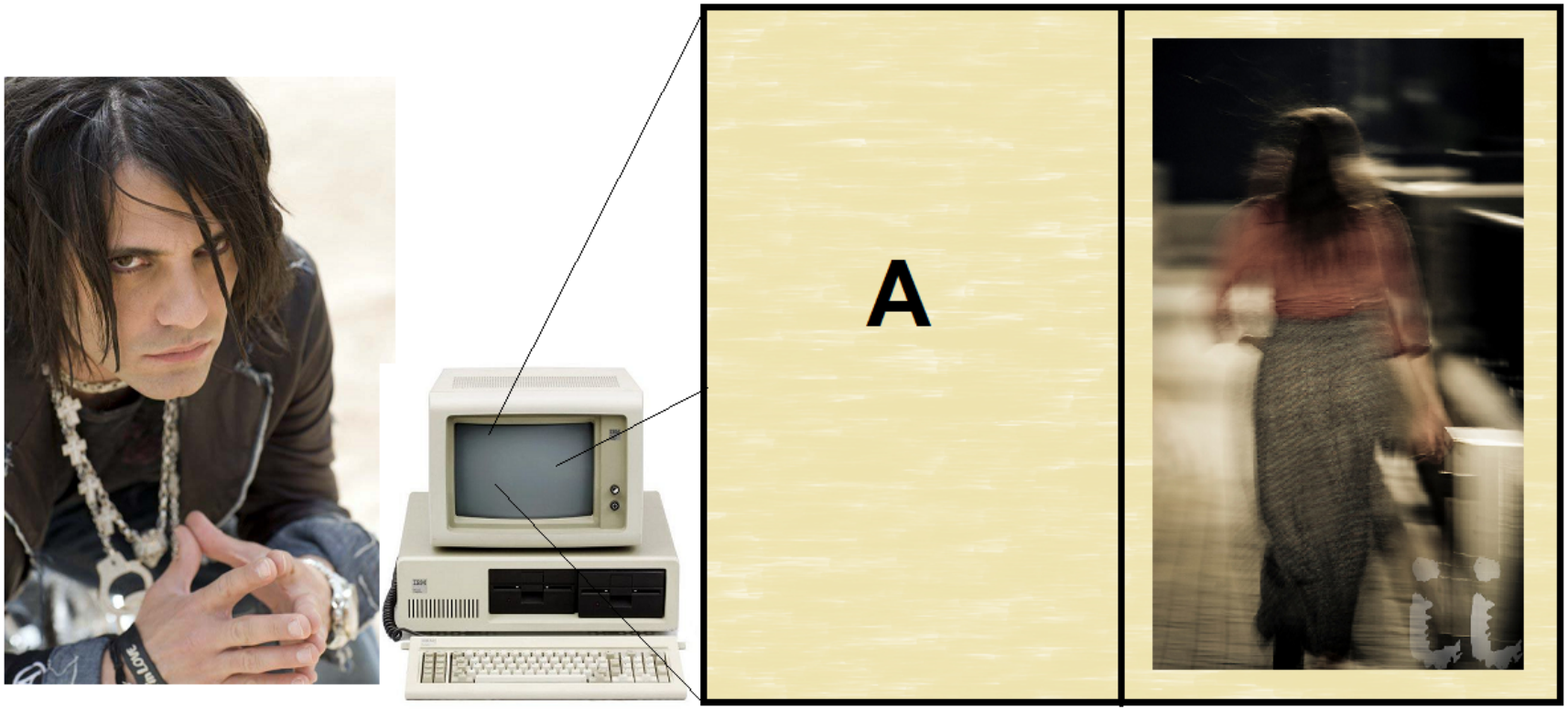




Experimental set-up



Experimental set-up



Experimental set-up

- Bem used images from the International Affective Picture System, a set of 820 pictures which are rated on 9-point scales for arousal level by men and women
- **The Hypothesis:** Participants will be able to identify the position of the hidden erotic picture significantly more often than by chance alone (50%).

- Only ($n =$)100 volunteers are recruited for this study, which is not very high.
- Furthermore, only **12** trials actually used erotic pictures, where the others used 'negative' pictures (maggots, dead animals) and 'neutral pictures' (apple).

  $\Rightarrow$ some problems are already starting to appear.
- Bem notes "I set 100 as the minimum number of participants [...] if $d = 0.25$ and $N = 100$, the power to detect an effect significant at 0.05 by a one-tail, one-sample t test is 0.80"

- Bem states that "participants correctly identified the future position of the erotic pictures *significantly more frequently* than the expected 50% hit rate expected", but their result was 53.1%. Is this really that significant, regardless of statistical significance? (Rule 5: Statistical Analysis Is More Than a Set of Computations)
- He later goes on to say that the 51.3% hit rate of negative pictures was not significant. In what context is a 3.1% difference from chance more significant than 1.3%?

- Bem also claims that the 53.1% hit rate is significant by a binomial test with $p = 0.11$. This is not what I found at all:

```
> binom.test(53, 100, 1/2)

        Exact binomial test

data:  53 and 100
number of successes = 53, number of trials = 100, p-value
= 0.6173
alternative hypothesis: true probability of success is not equal to
 0.5
95 percent confidence interval:
 0.4275815 0.6305948
sample estimates:
probability of success
                  0.53
```

- This is a good time to note that Bem does not include **a single confidence interval** in this entire paper.

- Later on, Bem claims that people who were deemed to be "stimulus seeking" individuals, as tested by the two questions "I am easily bored" and "I often enjoy seeing movies I've seen before", had a hit rate of 57.6%. Again with no confidence intervals or any discussion of the variation whatsoever.
- After describing the results of his 9 experiments, Bem does go on to discuss issues of replication, but these don't fully explain the issues with his study. His arguments seem to be more about *other psychologists and their beliefs*, rather than *actual statistical methodology*.

Conclusion of Issues:

- Small sample size ($N = 100$). In particular, a small number of participants were given the erotic images ($n = 16$).
- No discussion of variability which could indicate the (in)significance of the hit rate's difference from 50% – confidence intervals!
- The experiments were not replicable by other social scientists (and many attempted to do so, apparently).
- The binomial test I tried to reproduce was inconsistent with Bem's results (you need about $x \geq 61$ to get $p < 0.05$).

Bem ends the paper with the following remark:

Near the end of her encounter with the White Queen, Alice protests that "one can't believe impossible things," a sentiment with which the 34% of academic psychologists who consider psi to be impossible would surely agree. The White Queen famously retorted, "I daresay you haven't had much practice. When I was your age, I always did it for half-an-hour a day. Why, sometimes I've believed as many as six impossible things before breakfast"
(Alice in Wonderland – Carroll, 2006, p. 166).
Unlike the White Queen, I do not advocate believing impossible things. But perhaps this article will prompt the other 66% of academic psychologists to raise their posterior probabilities of believing at least one anomalous thing before breakfast.

Failing the Future: Three Unsuccessful Attempts to Replicate Bem's 'Retroactive Facilitation of Recall' Effect

Stuart J. Ritchie, Richard Wiseman, Christopher C. French

Published: March 14, 2012 • https://doi.org/10.1371/journal.pone.0033423

Precognition or Pathological Science? an Analysis of Daryl Bem's Controversial "Feeling the Future" Paper

By Gauvrit, Nicolas

Daryl Bem Proved ESP Is Real

Which means science is broken.