What Counts as Credible Evidence in Contemporary Evaluation Practice?

Stewart I. Donaldson, Ph.D.
Claremont Graduate University
stewart.donaldson@cgu.edu

Hawaii-Pacific Evaluation Association
Annual Conference
September 7, 2007

Hilton Waikiki Prince Kuhio Hotel

Overview

- Contemporary Evaluation Practice
- Evidence-based Practice
- What Counts as Credible Evidence?
- Discussion of the Gold Standard
- Practical Program Evaluation

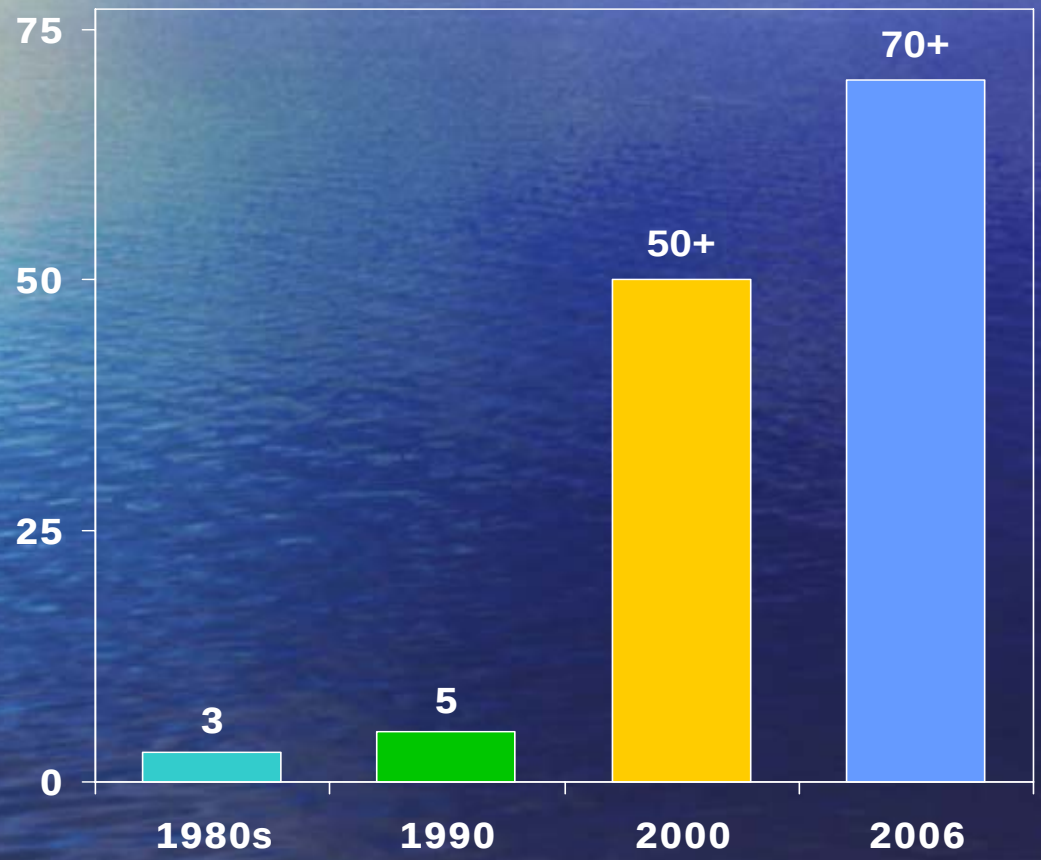
Contemporary Evaluation

- Evaluation Theory
- Evaluation Design
- Evaluation Methods
- Evaluation Practice
- The Evaluation Profession
- Research on Evaluation

An Indicator of the Second Boom in Evaluation Practice

- 1980s – Only 3 National and Regional Evaluation Societies
- 1990 – 5
- 2000 – More than 50
- 2006 – More than 70 including a Formal International Cooperation Network

Number of Evaluation Professional Associations



The Top Regional Evaluation Association of the Future

Hawaii-Pacific Evaluation Association

Evidence-Based Practice

- Highly Valued
- Global
- Multidisciplinary
- Many Applications

Sample of Applications

- Evidence-based Medicine
- Evidence-based Mental Health
- Evidence-based Management
- Evidence-based Decision Making
- Evidence-based Education
- Evidence-based Coaching

Sample of Applications

- Evidence-based Social Services
- Evidence-based Policing
- Evidence-based Conservation
- Evidence-based Dentistry
- Evidence-based Policy
- Evidence-based Thinking about Health Care

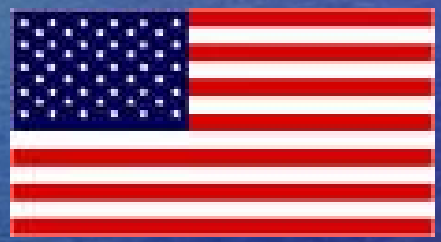
Sample of Applications

- Evidence-based Occupational Therapy
- Evidence-based Prevention Science
- Evidence-based Dermatology
- Evidence-based Gambling Treatment
- Evidence-based Sex Education
- Evidence-based Needle Exchange Programs
- Evidence-based Prices
- Evidence-based Education Help Desk

New Formula

Mom

+



The Flag

+

Warm Apple Pie

=

Evidence-based Practice

In God We Trust

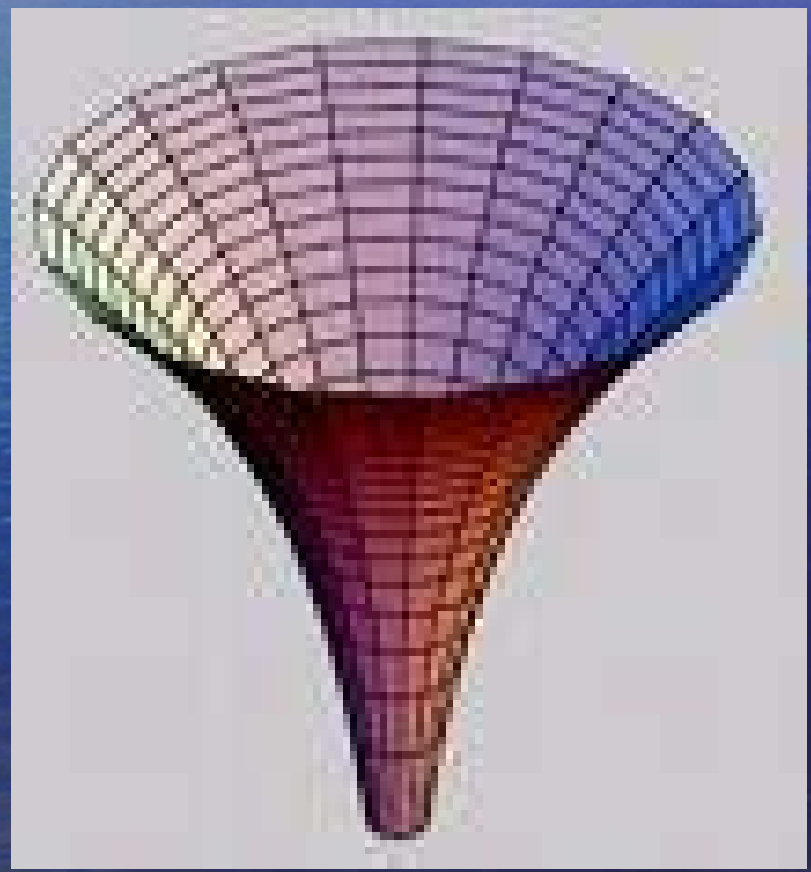
- ALL OTHERS MUST HAVE CREDIBLE EVIDENCE

A Popular Alternative to Evidence-based Practice

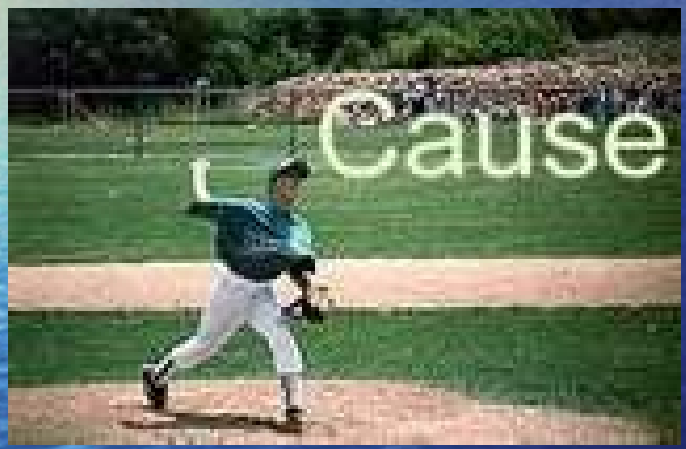
- FAITH-BASED PRACTICE & PROGRAMS

What Counts as Credible Evidence?

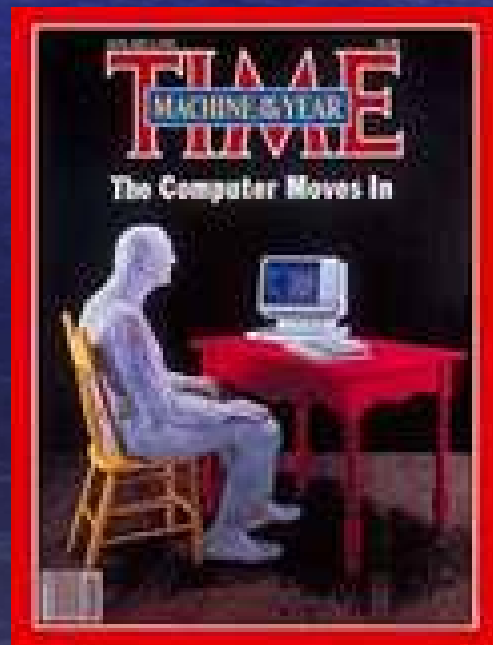
Wide Range of Views about Credible Evidence



The Cause-Effect Challenge in Evaluation

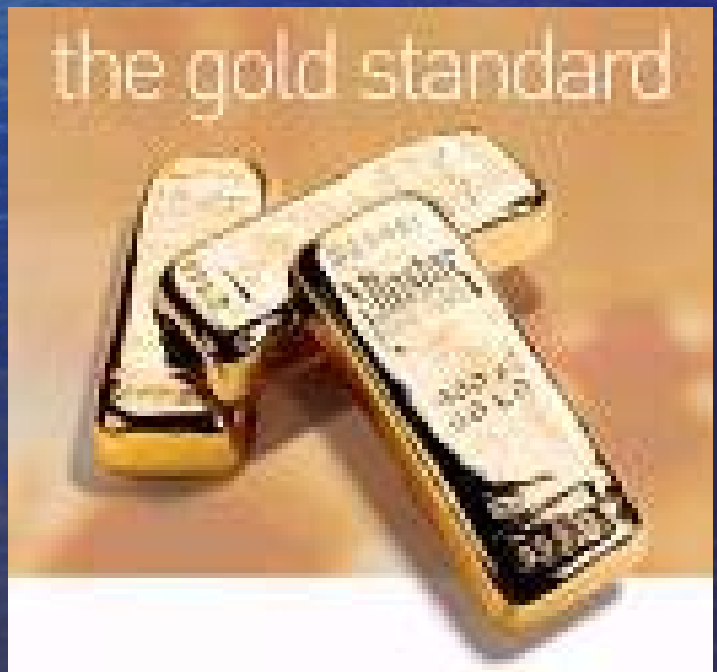


The Solution: In Search of a Time Machine



Experimental Design: Gold Standard?

- Random Assignment
- Experimental Control
- Ruling Out Threats to Validity



Supreme Courts of Credible Evidence

- What Works Clearinghouse
- Campbell Collaboration
- Cochrane Collaboration

Bias Control on Judgment Day

Evaluation Design Dominance

Single Rigor Hierarchy

Versus

Situational Variation and Appropriateness

(Patton, 2007)

Challenges of the Gold Standard

- AEA Statement vs. Not AEA Statement
- Theoretical
- Practical
- Methodological
- Ethical
- Ideological
- Political

AEA Statement vs. Not AEA Statement

- AEA Opposition to Priority on RCTs
  - "Privileging RCTs: Back to the Dark Ages"
  - "Priority Manifests Fundamental Misunderstandings Causality and Evaluation"
- AEA Members Opposition to AEA Statement
  - "Lack of Input from Key AEA Members"
  - "Unjustified, Speciously Argued, Does Not Represent Norms or Many AEA Members Views"
- AEA Compared to the Flat Earth Society

Diverse Prescriptive Theories of Evaluation Practice

- Social experimentation
- Science of valuing
- Results oriented management
- Utilization focused evaluation
- Empowerment evaluation
- Realist evaluation
- Theory-driven evaluation
- Inclusive evaluation
- Fourth generation evaluation

RCTs Not Practical/Feasible

- Often Impossible to Implement Well
- Not Cost Effective
- Very Limited Range of Applications
- Evidence to the Contrary

RCT Ethical Issues

- Unethical to Withhold Treatment from Control Groups
- Why Evaluate if Treatment is Better?
- Delay Treatment
- Non Evidence-Based Programs are Unethical

Methodological Challenges

- Zero Blind vs. Double Blind – Experimenter Effects
- Allegiance Effects
- Unmasked Assignment
- Misguided Arguments about Causality
- External Validity Concerns
- Recent Developments to Overcome Some Challenges noted in the Past

Political Concerns

- The RCT Gang has hijacked the term "evidence-based" for political and financial gain
- "Evidence" and especially "scientific or rigorous evidence" have become code for RCTs
- Focusing evaluation around these particular ideas about "scientific evidence," allows social inquiry to become a tool for institutional control and to advance policy in particular directions

Political Concerns

- It is epistemological politics, not the relative merits of RCTs, that underlie federal directives on methodology choice
- The demand for evidence advances a “master epistemology.” The very dangerous claim is that a single epistemology governs all science
- Privileging the interests of the elite in evaluation is radically undemocratic

Ideological Differences: Paradigm Wars

- “The positivist can’t believe their luck, they’ve lost all the arguments of the last 30 years and they’ve still won the war!”

- “The world view underlying the current demand for evidence is generously speaking a form of conservative post-positivism, but in many ways is more like a kind of neo-positivism.”

Ideological Differences: Paradigm Wars

- Many of us thought we'd seen the last of this obsolete way of thinking about the causes and meanings of human activity, as it was a consensual casualty of the great quantitative-qualitative debate in the latter part of the 20th century.
- Human action is not like activity in the physical world.
- Social knowledge is interpreted, contextual, dynamic or even transient, social or communal, and quite complicated. Privilege and honor complexity.

Ideological Differences: Paradigm Wars

- Evidence-based evaluation concentrates evaluation resources around one small question, does the program work?, and uses but one methodology, despite a considerable richness of options. The result is but one small answer.

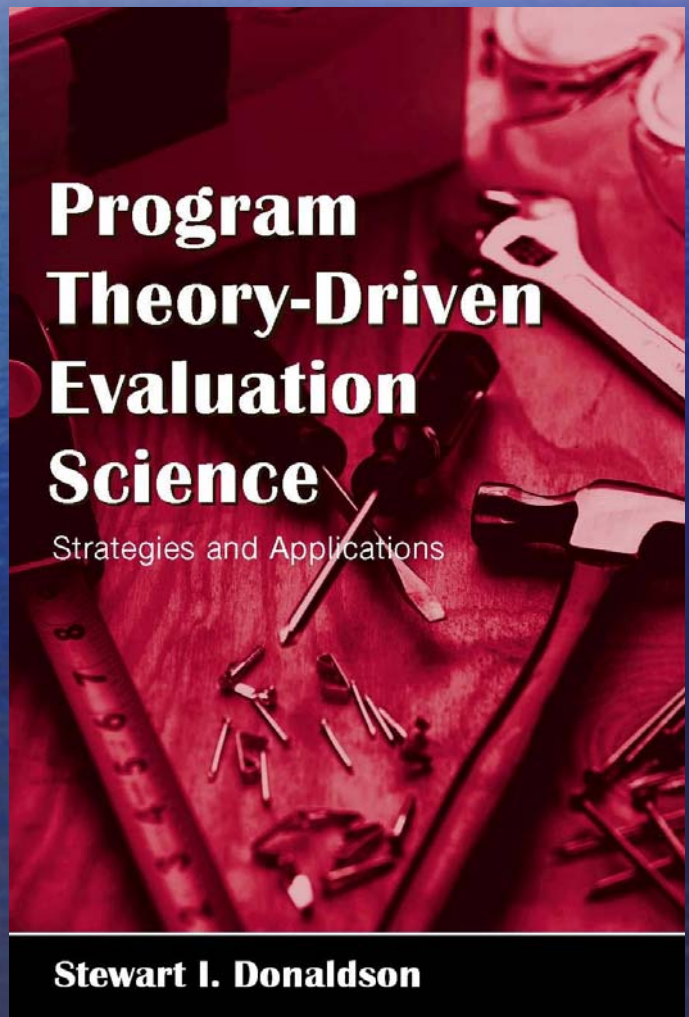
- So what kind of evidence is needed? Not evidence that claims purchase on the truth with but a small answer to a small question, neat and tidy as it may be.

So What Kind of Evidence is Needed? (Greene, in press)

Evidence:

- that provides a window into the messy complexity of human experience
- that accounts for history, culture, and context
- that respects differences in perspective and values
- about experience in addition to consequences
- about the responsibilities of government not just responsibilities of its citizens
- with the potential for democratic inclusion and legitimization of multiple voices - evidence not as proof but as inkling

Practical Program Evaluation: A Program Theory Approach



Practical Program Evaluation

- **Integrative Framework**
- **Contextual: Contingency Perspective**
- **Method Neutral**
- **Culturally Competent**
- **Evaluation Standards**
- **Guiding Principles**

Gather Credible Evidence

Definition: Compiling information that stakeholders perceive as trustworthy and relevant for answering their questions. Such evidence can be experimental or observational, qualitative or quantitative, or it can include a mixture of methods. Adequate data might be available and easily accessed, or it might need to be defined and new data collected. Whether a body of evidence is credible to stakeholders might depend on such factors as how the questions were posed, sources of information, conditions of data collection, reliability of measurement, validity of interpretations, and quality control procedures (CDC Evaluation Framework).

Gather Credible Evidence (Continued)

- Role: Enhances the evaluation's utility and accuracy; guides the scope and selection of information and gives priority to the most defensible information sources; promotes the collection of valid, reliable, and systematic information that is the foundation of any effective evaluation (CDC Evaluation Framework).

Some Guiding Principles for Evaluation Practice

- Understand the Diversity of Prescriptive Evaluation Theories
- Be Aware of the Diverse Views on what Counts as Credible Evidence
- Ongoing Discussions of Stakeholder Expectations about Evidence
- Secure Buy-in to the Evaluation Design Before Revealing Results
- Be Aware of Potential Standards of Judgment
- Be Prepared for Meta-Evaluation
- Credible Evidence is Often Key for Evaluation Influence & Positive Change

An Ounce of Credible Evidence is Worth a Ton of Enthusiasm and Good Intentions