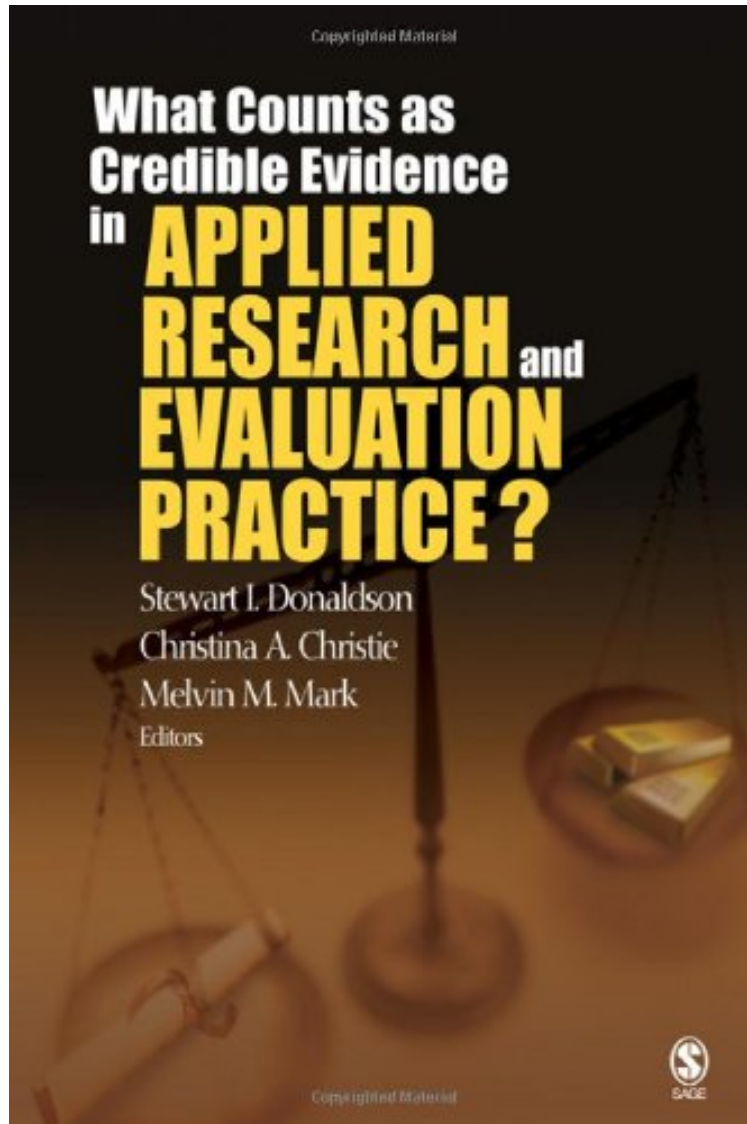


(Mobile book) What Counts as Credible Evidence in Applied Research and Evaluation Practice?

What Counts as Credible Evidence in Applied Research and Evaluation Practice?

Stewart I. Donaldson, Christina A. Christie, Melvin M Mark
DOC | *audiobook | ebooks | Download PDF | ePub



[Download](#)

[Read Online](#)

#1553910 in Books SAGE Publications, Inc 2008-10-22 Original language: English PDF # 1 .70 x 5.90 x 8.80l, .75 #File Name: 1412957079288 pages | File size: 69.Mb

Stewart I. Donaldson, Christina A. Christie, Melvin M Mark : What Counts as Credible Evidence in Applied Research and Evaluation Practice? before purchasing it in order to gauge whether or not it would be worth my time, and all praised What Counts as Credible Evidence in Applied Research and Evaluation Practice?:

2 of 2 people found the following review helpful. Applicability for PhD programs in Nursing By Tamah I find this book very helpful in researching arguments for and against random controlled trials as the only form of credible research.

There are several perspectives that catalyze further thought and research on the subject. This book is well worth the cost and time to read it. I believe its applicability is long-standing for other classes as well. 1 of 1 people found the following review helpful. Balanced text to bring some sense to evaluation paradigm wars

By J. I. Uitto

The evaluation field continues to engage in paradigm wars that involve heated debates over which approaches and methodologies produce the most reliable results to support evidence-based policy-making. Somewhat regrettably, the commendable goal of enhanced rigour in evaluation research has been hijacked by a focus on a narrow set of experimental methods--randomized controlled trials or RCTs--which have been proclaimed as the 'gold standard' by their proponents. This trend has been boosted by calls for unambiguous measurements of results and impacts, and cost-efficiency, by policy-makers and bureaucrats struggling with making policy choices and undertaking programs under increasing resource constraints. On the other side, the reaction from the proponents of more qualitative methodologies and participatory approaches to evaluation has been strong, even emotional at times. As a professional evaluator, I've witnessed these brawls first hand. This book makes an excellent contribution to the debate through a balanced presentation of the issues and by letting the different sides to make their respective cases. The authors in the book include a number of leading scholars and practitioners in the field. The perspective is North American (all authors work in the US or Canada) and draws heavily on experiences from education and social services. Although my own work pertains to evaluating international development programs, I found the discussion in the book on what constitutes credible evidence very valuable. In the two introductory chapters the editors frame the debate in the context of a search for evidence-based society and how this has played out in the use of experimental and non-experimental designs for collecting evidence. Quite didactically, they place the debate within the broader scientific paradigms of social inquiry, including logical positivism, post-positivism, constructivism and related thinking, and pragmatism. They also tentatively place the chapter authors along these paradigmatic axes. The following eight chapters are divided into two sections arguing, respectively, for experimental and non-experimental routes to credible evidence. The trajectory of the chapters moves from the most hard-core case for experimental designs to an argument for the credibility of image-based research in evaluation. The chapters in between tend to take a conciliatory approach to the extent that the last of the chapters in the experimental section could have been moved to the non-experimental section.

Part II entitled 'Experimental Approaches as a Route to Credible Evidence' contains four chapters. In the first, Gary Henry argues for high-quality policy and program impact evaluations as a necessity for providing solid evidence for policy choices, linking the matter to democratic theory and the need to detect and debunk bad policies. His view assumes that there actually are ways to objectively define what produces desirable outcomes--and what such outcomes are. He acknowledges types of bias in evaluation research, but sees them in technical, rather than political terms. In the chapter that follows, Leonard Bickman and Stephanie Reich assess the credibility, reliability and validity of RCTs, concluding that while there are threats to the validity (especially external validity) of RCTs, they still can be seen to be amongst the most credible designs available to evaluators. Despite this overall conclusion, Bickman and Reich acknowledge that there are other, non-experimental approaches to establish causality (many natural science disciplines--geology, astronomy, engineering, and subfields of medicine--base their research on non-experimental designs). In social sciences, they particular highlight the program theory, theory-driven or pattern-matching method, recognizing that such other approaches are needed to supplement RCTs that can only answer a very limited number of questions. The last chapter in this part of the book, by George Julnes and Debra Rog, introduces the concept of actionable evidence. The authors assert that for evidence to be useful, it should not only be credible but also actionable, defined as adequate and appropriate for guiding actions in targeted real-world contexts. The lengthy chapter takes as its starting point the question of relating the choice of methods to the questions that stakeholders want addressed. They proceed to outline a multitude of evaluation tasks (borrowing from Carol Weiss) and then consider the implications of the tasks for methodology. Another way of framing the evaluation questions relates to the level of conclusion and the different levels of causal questions in impact evaluation, including whether the evaluation seeks to provide an aggregate description, disaggregation for causal analysis, or an inferential analysis of the underlying constructs and causal mechanisms. They state that, "experimental methods are argued as appropriate for strengthening impact-evaluation conclusions, but the value of these methods is dependent on the level of conclusions being addressed" (p. 104). Summarizing the discussion on the relationships between questions and methods, Julnes and Rog express their view that, while particular questions call for quantitative designs, there is substantial territory open to other designs. This summary leads the authors to consider the contextual factors that affect the adequateness and appropriateness of alternative methods, including policy context and the nature of the phenomena studied. They then discuss how to judge the adequacy of methods for providing the evidence that is needed to address the stakeholder questions identified, and when is it appropriate to use particular methods for causal analysis, taking into account constraints posed by factors internal to the program, evaluation capacity and political constraints, as well as ethical considerations. The discussion is nuanced and fair. Julnes and Rog conclude by affirming the primacy of the evaluation stakeholder questions in influencing the types of evidence needed. They caution against "simple frameworks that drive method choice in a somewhat automatic fashion" and instead wish to support "more informed judgments on method choice in political public policy environments" (p. 128). To me, this is one of the strongest, most thoughtful and balanced

chapters in the book. Therefore, it also deserves its place in the middle, bridging the quantitative and qualitative parts. Another chapter in this section of the book, preceding the chapter discussed above, by Russell Gersten and John Hitchcock focuses on the role of the What Works Clearinghouse, established in 2002 by the U.S. Department of Education. The chapter is descriptive and possibly useful to education researchers and evaluators, but did not raise my own somewhat biased interest. Part III on 'Nonexperimental Approaches for Building Credible Evidence' consists of another five chapters, starting with one of the grand old men of evaluation, Michael Scriven. His chapter, 'Demythologizing Causation and Evidence,' is written with flair in lively language. Like Julnes and Rog before him, Scriven enlists other sciences--from mathematical physics and geology to anthropology, ethnography and epidemiology--to demonstrate how experimental methods are but one of the many approaches to analyse causation. He writes, "much of the world of science, suffused with causal claims, runs along very well with the usual high standards of evidence, but without RCTs" (p. 136). Breezing through the origins of causal concepts, the cognitive process of causal inference vs. observation, and the level of evidential certainty required for scientific, legal and practical purposes, he then addresses the alleged supremacy of RCTs as well as "other contenders." His myth-busting position is that "(i) the attempted takeover of the terms evidence and cause is partly inspired by the false dichotomy between experiment and quasi-experiment, and (ii) the whole effort is closely analogous to the attempted annexation of the concept of significance by statistically significant" (p. 151). The following chapter by Jennifer Greene discusses evidence as 'proof' vs. evidence as 'inkling.' Her premise is that evaluation is both influenced by the political, organizational and sociocultural contexts, as well as it serves to shape that context. Consequently, "evaluation is not a bystander or neutral player in the debates that often surround it, but rather an active contributor to those debates and to the institutions that house them" (p. 153). Her chapter attempts to demonstrate how the present discourse assumes that 'evidence' can make social systems 'efficient and effective' and how these assumptions convey a particular view of human phenomena and responsibilities of government in democratic societies. The argument is a useful antidote to the positivistic view presented by Henry earlier in the book. Her vision of evidence is not providing the truth, or neat and tidy small answers to small questions. Rather, in Greene's view evidence must provide a "window into the messy complexity of human experience; evidence that accounts for history, culture, and context; evidence that respects difference in perspective and values; evidence about experiences in addition to consequences; evidence about the responsibilities of government, not just the responsibilities of its citizens; evidence with the potential for democratic inclusion and legitimization of multiple voices--evidence not as proof but as inkling" (p. 166). Sharon Rallis starts her chapter by telling the story of how she first encountered evaluation when she was teaching a federally funded summer program that was subject to an evaluation. The evaluators insisted on holding on to their plan to assess the program against a single outcome, with no regard to the important associated benefits that the program had bolstering the self-esteem of the participating students. Furthermore, the evaluation based on a quasi-experimental design deprived half of the students from participating in an important part of the program, which Rallis and other program colleagues felt was unfair. While the evaluators claimed that their work was scientific and rigorous, Rallis pondered about the missing piece and came to the conclusion that it was 'probity'--goodness and moral soundness. Consequently, she began to study evaluation with a commitment to make evaluations useful for the program personnel and participants. This chapter elaborates on her vision of evaluation with probity and moral reasoning, grounded in nonconsequentialist theories. She explains: "The evidence we collect looks quite different from that of our colleagues who measure outcomes. Our aim is not to cast judgment ... but to discover what happened and what the experience meant to the program participants. We hope that our discoveries can lead to improving the program and thus the well-being of the participants" (pp. 174-175). Rather than RCTs, evaluation done with these principles borrows tools from fields such as ethnography, phenomenology and sociolinguistics/semiotics. She presents a case from an evaluation and needs assessment of an HIV/AIDS education and prevention program that provides some unexpected insights into the participants' experiences. She asserts that this work is rigorous because "it is grounded in theory and previous research and in moral principles of justice and caring" (p. 178). Sandra Mathison in her chapter 'Seeing Is Believing' explores the credibility of image-based research and evaluation, as one form of evidence to establish and represent truth and value. Like the part III authors before her, she emphasizes how the credibility of evidence and the knowledge thus created is contingent on experiences, perception and social conventions. Image-based research uses images in three ways: (i) as data or evidence; (ii) as an elicitation device to collect other data, and (iii) as a representation of knowledge (p. 184). Mathison posits four considerations for establishing the credibility of image-based research: (1) quality of the research design, (2) attention to context, (3) adequacy of the image from multiple perspectives, and (4) the contribution images make to new knowledge (p. 188). The last chapter in part III, by Thomas Schwandt, is entitled 'Toward a Practical Theory of Evidence for Evaluation' and it functions as a kind of recap of what has come before; as such, it could have equally well been placed in part IV on conclusions. This is another rich chapter that goes to the heart of the debate of what we mean by evidence: "...information helpful in forming a conclusion or judgment. Framed in a more rigorous epistemological perspective, evidence means information bearing on whether a belief or proposition is true or false, valid or invalid, warranted or unsupported. At present, we face some difficulty and confusion with understanding the term evidence in evaluation because it is often taken to be synonymous with the term evidence-

based" (p. 199). He then proceeds to problematize the term evidence-based, as being narrowly interpreted to mean only a specific kind of finding regarding causal efficacy. Secondly, Schwandt argues why evidence cannot serve as a secure and infallible base or foundation for action. Furthermore, he emphasizes that, as an aspect of policy making, evaluation must consider ethics. Schwandt concludes that "deciding the question of what constitutes credible evidence is not the same as deciding the question of what constitutes credible evaluation ... However necessary, developing credible evidence in evaluation is not sufficient for establishing the credibility of an evaluation" (p. 209). He further asserts that method choice alone does not determine what is credible and convincing evidence. He calls for framing evidence in a practical-theoretical way that is concerned with the character and ethics of evidence and the contexts in which evidence is used. In the final part of the book, Melvin Mark summarizes the different perspectives of the book with an aim of changing the terms of the debate. He concludes: "Extensive and continued discussion of the relative merits and credibility of RCTs versus other methods would have limited capacity to move forward our understanding and our practice. ... by changing the terms of the debate, we may be able to improve understandings of deeply entrenched disagreements; move toward a common ground where such can be found; better understand the disagreements that remain; allow, at least in select places, credible evidence as part of the conversation; and enhance the capacity of stakeholders to make sensible decisions rather than be bewildered by our disagreement or draw allegiances based on superficial considerations" (pp. 237-8). Certainly a deserving goal. The book ends with an epilogue by Stewart Donaldson that attempts to provide a practitioners guide for gathering credible evidence in the evidence-based global society. I thoroughly enjoyed reading this book and although much of the debate around epistemology, approaches and methods is familiar to someone educated and working in evaluation and applied social science research, the way it is framed in this book is truly enlightening. For a thoughtful reader, it becomes evident that the truth--as almost always in such debates--is somewhere in the middle. All of the approaches and methodologies have merit when used appropriately and in appropriate contexts. Both experimental and non-experimental methods can be rigorous, but both can also have serious flaws with regard to internal and external validity, relevance and appropriateness. The old saw about everything looking like a nail when you have only a hammer in your tool box is true here as well. The take-home lesson is that, instead of allowing methods to dictate one's evaluation questions and designs, one should choose one's methods according to the questions one wants answered.

"The issue of what constitutes credible evidence isn't about to get resolved. And it isn't going away. This book explains why. The diverse perspectives presented are balanced, insightful, and critical for making up one's own mind about what counts as credible evidence. And, in the end, everyone must take a position. You simply can't engage in or use research and evaluation without deciding what counts as credible evidence. So read this book carefully, take a position, and enter the fray." Michael Quinn Patton, Author of *Utilization-Focused Evaluation*, 4e "I found this text to be very interesting and useful in capturing and presenting varying perspectives in the field. There are some very good points and considerations for students and practitioners in this book." Michael Schooley, Centers for Disease Control "Donaldson and colleagues have assembled an insightful and timely collection of papers on the complex issues regarding what constitutes credible evidence in evaluation. This important book offers readers the latest thinking on generating actionable evidence for policy and program decision-making from a wide variety of philosophical perspectives. The book is an indispensable resource for evaluation scholars and practitioners on this longstanding and central debate in the evaluation field." Robin Lin Miller, Michigan State University Placing into perspective the meaning of evidence for evaluation professionals and applied researchers, this text provides observations about the diversity and changing nature of credible evidence, Editors Stewart I. Donaldson, Christina A. Christie, and Melvin M. Mark include lessons from their own applied research and evaluation practice, and suggest ways in which practitioners might address the key issues and challenges of collecting credible evidence. Key Features Provides summaries of the strengths and weaknesses of the varied approaches to research and evaluation to give readers greater insight and guidance on how to select the appropriate methods for their work Offers diverse definitions of "evidence" so that readers can evaluate the landscape of this highly debated research issue Devotes a full chapter to the implications of evidence for contemporary applied research and evaluation practice This book is appropriate for a wide range of courses, including Introduction to Evaluation Research, Research Methods, Evaluation Practice, Program Evaluation, Program Development and Evaluation, and evaluation courses in Social Work, Education, Public Health, and Public Policy.

"[The book's] design is a good one, containing papers for, papers against, and papers offering a synthesis and summing up." -- Kenneth Watson * *The Canadian Journal of Program Evaluation* * "[The books] design is a good one, containing papers for, papers against, and papers offering a synthesis and summing up." (Kenneth Watson *The Canadian Journal of Program Evaluation* 2011-01-01) About the Author Stewart I. Donaldson is Professor and Director of the Claremont Evaluation Center, and Dean of the Schools of Social Science, Policy Evaluation and Community Global Health at Claremont Graduate University. Dean Donaldson continues to develop and lead one of the most extensive and rigorous graduate programs specializing in evaluation. Dr. Donaldson is currently serving as the

Director of the American Evaluation Associations (AEA) Graduate Education Diversity Internship (GEDI) Program and recently served a 3-year term on the AEA Board. He leads the Certificate for the Advanced Study of Evaluation Program at Claremont (a distance education program for working professionals) and has taught thousands of graduate students and working professionals participating in online courses, workshops, webinars, and various other e-learning experiences. He is a fellow of the Western Psychological Association, serves on the Boards of the International Positive Psychology Association and EvalPartners, and the Editorial Boards of the American Journal of Evaluation, New Directions for Evaluation, Evaluation and Program Planning, and the Journal of Multidisciplinary Evaluation. Professor Donaldson has authored or co-authored more than 200 evaluation reports, scientific journal articles, and chapters and his recent books include *Credible and actionable evidence: The foundation of rigorous and influential evaluations* (this volume); *Practical program design and redesign: A theory-driven approach to program development and developmental evaluation* (forthcoming); *Evaluation for an equitable society* (forthcoming); *Theory-driven positive psychology: A culturally responsive scientific approach* (forthcoming); *Emerging Practices in International Development Evaluation* (2013); *The Future of Evaluation in Society: A Tribute to Michael Scriven* (2013); *Teaching Psychology Online: Tips and Strategies for Success* (2012); *Social Psychology and Evaluation* (2011); *Advancing Validity in Outcome Evaluation: Theory and Practice* (2011); *Applied Positive Psychology: Improving Everyday Life, Health, Schools, Work, and Society* (2011); *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* (2008); *Program Theory-Driven Evaluation Science: Strategies and Applications* (2007); *Applied Psychology: New Frontiers and Rewarding Careers* (2006); and *Evaluating Social Programs and Problems: Visions for the New Millennium* (2003). Dr. Donaldson has been honored with Early Career Achievement Awards from the Western Psychological Association and the American Evaluation Association. In 2013, he was honored with the American Evaluation Associations Paul F. Lazarsfeld Award for sustained lifetime written contributions to advancing evaluation theory and practice, and was elected President of the American Evaluation Association.

Christina A. Christie is a Professor and Head of the Social Research Methodology Division in the Graduate School of Education and Information Studies at University of California, Los Angeles. Christie specializes in educational and social policy and program evaluation. Her research focuses on the factors and conditions that influence evaluation practice in an effort to strengthen our understanding of evaluation as a method for facilitating social change. She has published extensively and her work appears in journals such as *American Journal of Evaluation*, *Children and Youth Services*, *Evaluation and Program Planning*, *Studies in Educational Evaluation* and *Teachers College Record*. Christie has served on the board of the American Evaluation Association (AEA) and is the former Chair of the Theories of Evaluation Division and the Research on Evaluation Division of AEA. Currently, she is an Associate Editor for the *American Journal of Evaluation*.

Melvin M. Mark is professor of psychology at the Pennsylvania State University, where he also is Head of the Department of Psychology. He has served as President of the American Evaluation Association (AEA). He was Editor of the *American Journal of Evaluation* (and is now Editor Emeritus). A social psychologist, Dr. Mark has wide ranging interests related to the theory, methodology and practice of evaluation, as well as a general interest in the application of social psychology to evaluation and applied social research. Dr. Marks awards include the American Evaluation Associations Lazarsfeld Award for Contributions to Evaluation Theory. He is author of more than 125 articles and chapters in books. Among his books are *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs* and the co-edited volumes *Social Science and Social Policy*; *SAGE Handbook of Evaluation*; *What Counts as Credible Evidence in Applied Research and Evaluation Practice*; *Evaluation in Action: Interviews with Expert Evaluators*; and *Social Psychology and Evaluation*.