Phillips, D. C. (2007). Adding complexity: Philosophical perspectives on the relation-
    ship between evidence and policy. In P. A. Moss (Ed.), *Evidence and decision
    making: The 106th Yearbook for the National Society for the Study of Education*
    (pp. 376–402). Malden, MA: Blackwell.
Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
Rossman, G. B., & Rallis, S. F. (2003). *Learning in the field*. Thousand Oaks, CA:
    Sage Publications.
Slater, J. K., & Constantine, N. (2003, September). *Sacramento County HIV/AIDS
    Prevention and Education Program needs assessment: What HIV+ persons tell
    us about prevention for positives*. Sacramento, CA: Center for Research on
    Adolescent Health and Development, Public Health Institute.
Strike, K., Haller, E., & Soltis, J. (1998). *The ethics of school administration*. New
    York: Teachers College Press.
Watson, J. D. (1968). *The double helix: A personal account of the discovery of the
    structure of DNA*. New York: New American Library.
Weiss, C. H. (1998). *Evaluation* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

10

Seeing Is Believing

The Credibility of Image-Based Research and Evaluation

Sandra Mathison

As for a picture, if it isn't worth a thousand words, to hell with it.

—Ad Rheinhardt, minimalist American painter

This chapter explores the credibility of image-based research and evalua-
tion, one form of evidence used to establish and represent truth and
value. While there are unique features of images and image-based inquiry,
their credibility exists within a larger framework for establishing the believ-
ability of facts and values.

The credibility of evidence, the knowledge it generates, is contingent on
experience, perception, and social conventions. As such, it can and does
change over time (consider that Pluto once existed as a planet, and now it
doesn't) and can and should be arrived at through many means, eschewing
the dominance of any one method. The credibility of evidence and the gen-
eration of knowledge is enhanced by embracing Feyerabend's (1975) notion
of an anarchist epistemology, the notion that every idea or strategy, however
new or old or absurd, may improve our knowledge of the social world.

Credible evidence cannot therefore be the province of only certain methods (such as observations or experiments) and cannot be expressed in only one way (such as statistical averages or vignettes). To extend Feyerabend's notion, credible evidence should be humanitarian, and include embracing political ideology as a source of credibility since such ideologies may be important in overcoming the chauvinism of particular perspectives, especially ones that maintain the status quo.

At any given time, our knowledge is the best it can be and so a common-sense realism prevails—we are reasonably certain at a given time that this or that is true, and we can act on that knowledge with relatively high degrees of confidence. In some senses, certainty is based on cultural norms, or inter-subjective meanings, those things we collectively know and act on even if that knowledge is tacit. But because knowledge is contingent, there are limits to its certainty, and what we know will inevitably change. In general, the issue of certainty turns on sustaining a Cartesian dualism of a mind and a separate real world—rejecting this dualism refocuses the issue of certainty, away from mapping mind onto a physical reality and toward usable and sensible knowledge.

Qualities of good evidence include relevance, coherence, verisimilitude, justifiability, and contextuality. Contextuality is important in two respects. First is the contextuality of the evidence itself, as the charge that something is taken out of context is always a serious challenge to the credibility of evidence. So, context is paramount in establishing the credibility of evidence. Knowing the point of view (for example, economic, aesthetic, or political), who generates evidence, and how, are critical aspects of credibility. The more context provided for evidence, the more credible it is. Context is important in a second sense of providing specific criteria for judging the credibility of evidence. The most obvious example of this is culturally different criteria for establishing credibility—one might contrast Western scientific notions of credibility with aboriginal notions of what makes evidence credible. Adopting the anarchist position offered by Feyerabend allows for open-ended possibilities for what can count as credible evidence and therefore knowledge of the social world.

Seeing Is Believing, or Is It?

Images are all around us; we are all image makers and image readers. Images are a rich source of data for understanding the social world and for representing our knowledge of that social world. Image-based research has a long history in cultural anthropology and sociology as well as the natural

sciences, but is nonetheless still relatively uncommon (Collier & Collier, 1986). Images should not be romanticized, but neither should their value as data and knowledge be ridiculed or avoided. Especially in Western industrialized cultures, images are often associated with artistic expression, entertainment, and persuasion. Images are seen as intuitive, representing implicit and subjective knowledge, while numeric, text, and verbal data are more associated with fact, reason, and objective knowledge. Images, in fact, are no more suspect than any other sort of data, such as numbers or text. Images, like any data, can be used to lie, to question, to imagine, to critique, to theorize, to mislead, to flatter, to hurt, to unite, to narrate, to explain, to teach, to represent. This chapter will discuss what image-based research is what it can contribute to social research, and more particularly discuss its credibility.

We do not necessarily need to see to believe, and indeed we believe many things we cannot or do not see directly. The Humean skepticism about the knowability of causation is a key example of knowing something without being able to see it. The classic billiard ball example illustrates this skepticism—in the Humean perspective we cannot see causation but we infer it from seeing one billiard ball moving across the table, hitting another, and causing that second to move. Even when what we know is not something we directly see, it is based on seeing, what is referred to in scientific parlance as observation. Seeing is intricately connected with believing, and thus knowing at the individual, cultural, or global levels of knowledge. Indeed there is biological evidence that seeing, at least in the context of space, is more reliable than other sources of data, and that the central nervous system summarizes visual information in a statistically optimal way (Witten & Knudsen, 2005).

What Is Image-Based Research and Evaluation?

Image-based research includes found and researcher- or participant-produced videos, photographs, drawings, cartoons, and other visual forms of expression and representation. In anthropology, sociology, and psychology, images have been used for some time as data and as an alternative approach to representing research results because they offer a different form for researchers and participants to express their experience and present themselves and their knowledge.

> Images are essential to human sense-making. We see and think and communicate using images. Like words, images can be used, construed, and read in different ways and can serve multiple functions. . . . Like words, images are part of who

we are, who we think we are, and who we become—they are integral to questions of identity and purpose. Like other aspects of sense-making, how images create meaning is a dynamic process involving dialectical negotiation or interaction between the social and the personal aspects in any give culture. (Weber, n.d.)

Image-based research involves those things that are intrinsically visual (like place, clothing, art) but also includes the visual, both real and metaphoric, in all contexts. Examples of images are photographs, video/film, drawings, cartoons, graphs and charts, typography, graphic art, graffiti, computer graphics, television, architecture, signs, as well as what is seen in the mind's eye. This last sort of image may not be recognized as truly visual, but indeed we give meaning to and judge the adequacy of people, events, and products by the images we hold only in our mind. Images of failing schools, successful classrooms, urban blight, and rural communities reside in our mind even if we have never been to those places.

Images are used in three ways: (1) as data or evidence, (2) as an elicitation device to collect other data, and (3) as a representation of knowledge. In the first case, the images are themselves the data or evidence in an evaluation or research study. The classic example is Margaret Mead and Gregory Bateson's photographic study of Balinese life (Mead & Bateson, 1942). This book includes 759 still photographs with captions including the context of the photos, complemented by theoretical discussion. Mead and Bateson were motivated to move beyond what Mead described as "ordinary English words" because of their inadequacy for capturing and communicating the emotions of the South Sea Islanders; photography and film became their alternate grammar. Children's drawings are another example of images as data. In my investigations of high-stakes testing in schools, children's experiences of testing have been captured in drawings of themselves, the testing context, and their relationships with their teacher and classmates. Figure 10.1 shows a cartoon and Figure 10.2 a self-portrait by a fourth grader taking the state-mandated English Language Arts test.
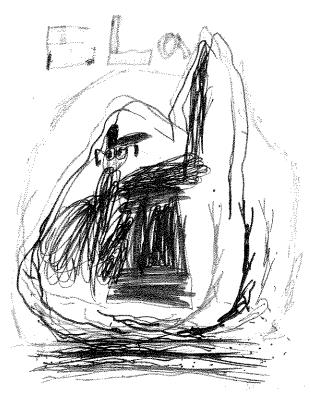
Using images to elicit other data, like cultural or personal meaning, is perhaps the most common form of image-based research (Harper, 2002). Auto-driven photography is a good example (see, for example, Clark, 1999). Evaluation or research study participants are given the opportunity to take photographs related to the inquiry questions at hand. The specific subject matter is left to participants, and as such the photographs become the physical manifestation of what participants pay attention to, perhaps what they value. The photographs are not the end goal, however. The meaning the photographer-participants make of the photographs, expressed in interviews, is the key data being collected. In an assessment of customer satisfaction with a hotel, Pullman and Robson (2007) asked guests to take photographs of things

Figure 10.1   Cartoon Drawing as Data

SOURCE: Freeman and Mathison (2005).

that influenced their opinions of the hotel. The photographs became the basis for discussions with guests about what they did and did not like.

This same strategy is used with photographs or video not generated by participants. Researchers who bring a theoretical framework to the research may take photographs or videos that focus the elicitation on specific research

Josef (Hispanic urban fourth grader): "[I was] worried 'cause if I don't, if I fail again 'cause I did third grade, 'cause in third grade . . . my teacher said 'some of you going to pass into fourth grade,' and I didn't . . . 'cause if I fail then I have to go to summer school. And I don't want to go to summer school again."

Figure 10.2   Joseph's Self-Portrait as Data

SOURCE: Freeman and Mathison (2005).

questions. For example, in a study of a therapeutic camp for children with cancer, the researchers hypothesized that the nature of physical space was a critical dimension of the camp's therapeutic quality. The photos of places and spaces at the camp were taken by the researchers and used in the interview with campgoers to explore this hypothesis (Epstein, Stevens, McKeever, & Baruchel, 2006). Regardless of who makes the images, photographs, or video recordings, they are used to elicit participants' thoughts, decisions, reasoning, and reactions in situ (see, for example, Meijer, Zanting, & Verloop, 2002).

Visual images, much like verbatim quotes, figures, graphs, and tables, can be used to illustrate the interpretive text—a legitimate, but particular use of images. The third use of images is as the representation of knowledge generated by a research study or evaluation. In other words, the image is the result of the research and is the communicative device for reporting findings. This use of images in research and evaluation challenges a taken-for-granted assumption that legitimate knowledge of what is or is not valued is best expressed in words, whether spoken or written. And the dominance of print media in academia reinforces this assumption. However, media other than text has become increasingly accessible and communicable with increased access to the World Wide Web. Large databases of images are now easily accessible, such as in the Landscapes of Capital Project (http://it.stlawu.edu/~global), an analysis of the role of corporate television commercials in shaping and defining global capitalism, and in Kaplan and Howes's (2004) Web-based compilation of photographs taken by students and staff in a secondary school.

The Kids With Cameras projects are another example of photographs as representation, made popular by the movie *Born Into Brothels*. While this project promotes photography as an empowering skill that will benefit these children, their photographs stand as representations of their life, a representation that needs no further interpretation or embellishment. The idea is that participant photography is a research strategy that permits participants to speak directly, and thus be empowered to influence their community, as well as decision and policy makers. The generic strategy is called "photovoice" (Wang, 1999; Wang & Burris, 1997). Photo essays are another example of images as representation, such as the representation of growing up in Los Angeles in Lauren Greenfield's (n.d.) *Fast Forward*, a series of photographs with contextualizing text. Peter Menzel's (1995) *Material World: A Global Family Portrait* (photographs taken by the photographer) and Jim Hubbard's (n.d.) *Shooting Back* (photographs taken by homeless children) are other excellent examples of image as knowledge.

Many graphs and charts are images, often communicating large data sets in simple or intuitive ways—Edward Tufte's work is the most significant in this area. Combining an aesthetic sensibility with a goal to communicate information, Tufte illustrates how descriptive information can be visualized to tell stories of multidimensional relationships. Tufte (1990) describes this as envisioning information, work he suggests is at "the intersection of image, word, number, art" (p. 9).

Credibility of Image-Based Research and Evaluation

All data are susceptible to manipulation, distortion, or misuse, but images no more so than any other kind of data. Just as statistics can be used in ways that distort meaning, so too can images distort meaning. "Photoshopped" has become a verb indicating an image has been digitally changed to deliberately

distort meaning, for example, and has been extended to other uses such as when the media focuses on celebrity breakups rather than the victims of Katrina or when history textbooks do not include embarrassing moments of racism or sexism. The more technical term for photographic manipulation is "scitexing" (named for the company that developed the technology) and can include overlaying multiple negatives, airbrushing, image quilting, and so on. There is a long history of the manipulation of images for whimsical, commercial, and political purposes, and credibility in image-based research must necessarily consider such possibilities (see Loizos, 2000, and the website *Is Seeing Believing?* at www.frankwbaker.com/isb.htm for more on the manipulation of images). But, researchers and evaluators are no more likely to manipulate image data than they are to manipulate statistical data.

Four considerations for establishing the credibility of image-based research are discussed below: (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.

Research Design Provides the Framework for Credibility

In part, the credibility of any data can be judged by reference to the design of the evaluation or research project (Wagner, 2004). There is a tremendous debate about what constitutes the best research design, marked by a recent advocacy for randomized clinical trials (and indeed, that is part of the motivation for this book). There is not, however, unanimity about what is best, in part because best for what and in what context is an important part of the equation. Without indulging in a discussion of the very different perspectives about what constitutes an adequate research design, the credibility of images can be established within the context of whatever research design is used.

Image-based research can operate within a neopositivist framework where image data are like any other kind of data. The credibility of the evidence, and the threats to credibility, are established procedurally. Research and evaluation studies intend to answer questions or test hypotheses. Whether the evidence collected provides credible answers or conclusions depends on such strategies as establishing the validity of the measures (i.e., are the images actually what is of interest), the reliability of images (i.e., do individuals or researchers agree on images), the sampling of images from a knowable population of images in an explicit and defensible way, and whether the analysis follows understood conventions.

An interpretivist or critical research or evaluation study that employs images as evidence will also make reference to procedures to establish credibility. For example, it is likely that participant-generated images would be favored over researcher-generated images, to be coherent with the intention to understand the meaningfulness of a construct, experience, program, and so on from the participants' point of view. Another example would be a clear and transparent justification for the selection of images to use in elicitation strategies. Yet another example would be a description of how analytic categories are generated and used in the analysis and conclusions. In addition, other conventions such as prolonged engagement (images collected over time, in multiple contexts) and responsiveness (the use of images reflects a respectful interaction between researcher/evaluator and participants) might be the focus.

An especially unique attribute of an interpretivist research design is the inclusion of a personal account of how and why the study was done. Personal accounts may be included in appendices, forewords, or afterwords, and outline the positionality of the study author, making explicit that persons and not merely technique shape the inquiry.

Regardless of the underlying epistemologies, the credibility of image-based research can be judged by the defensibility of the study design. Truth or credibility is, at least in part, established through conventions. While different camps of inquiry accept different conventions for establishing truth, there is a set of conventions to which a particular image-based evaluation or research study or image-based evidence can be compared. Image-based evaluation and research, like any, can therefore be established by reference to procedures that by agreement of a community of inquirers have come to be acceptable.

Credibility of Images as Evidence Requires Attention to Context

When they photographed social life in Bali, Bateson took pictures while Mead noted when, where, and under what circumstances the photograph was taken (Mead & Bateson, 1942). In presenting the images in their book, careful attention was paid to sequencing the photographs in the order they were taken. This was done in some cases because the photographs illustrate an event that has a temporal quality (a dance, a ceremony, preparation of a meal), but in a larger sense Mead and Bateson were providing a context within which they and the viewer/reader could understand Balinese social life. Visual sociology and other image-based research often use and make images that resemble other uses of images, such as photojournalism or art.

Visual sociology, documentary photography, and photojournalism, then, are whatever they have come to mean, or been made to mean, in their daily use in worlds of photographic work. They are social constructions, pure and simple. In this they resemble all the other ways of reporting what we know, or think we have found out, about the societies we live in, such ways as ethnographic reports, statistical summaries, maps, and so on. (Becker, 1998)

In his analysis, Becker (1998) illustrates that knowing who is creating and using images is a critical aspect of establishing their credibility. He asserts that images, like all cultural objects, get meaning from their context. A common means for providing context is the combination of images with text. Using Mead and Bateson's (1942) book on Balinese life as an example, when opened the pages facing each other provide a complement of image and accompanying contextualizing text: photographs appear on one page and two kinds of text on the other—a few paragraphs of interpretation and captions indicating when, where, and by whom the photograph was taken and what was going on at the time, as well as interpretive text suggesting the generalizable or theoretical meaning of the photographic evidence. Detail, then, is one way to provide context: "[I]n visual reasoning, art, typography, cartography, even sculpture, the quantity of detail is an issue completely separate from the difficulty of reading. Indeed, at times, the more intense the detail, the greater the clarity and understanding—because meaning and reasoning are contextual" (Tufte, 2003, p. 12). While talking about data displays in general, Tufte's (2006) emphasis on detailed documentation as the key to credibility applies to images of all types (p. 132).

Mead and Bateson's (1942) work exemplifies explicit, detailed contextualizing of images, but credible image-based research may also imply the context. "Because the images themselves, sequenced, repetitive, variations on a set of themes, provide their own context, they teach viewers what they need to know in order to arrive, by their own reasoning, at some conclusions about what they are looking at" (Becker, 1998, p. 14). The credibility of images is enhanced by providing or implying context (whether through text, narration, or other images) and by using complex and detailed images.

Adequacy of the Interpretation of the Image

The adequacy of the interpretation of images as data is the third aspect of credibility, and can be understood in terms of the focus on the subject matter, focus on the creation, or focus on the audience or viewers of the images (Werner, 2006). One or the other of these foci may be more relevant in establishing the credibility in any particular study, but regardless of the foci, at the heart of the adequacy of the interpretation is a sense of authenticity. Table 10.1 summarizes these foci, including key questions one might ask to establish the credibility of an image or an image-based study.

In assessing the adequacy of the interpretations in image-based studies, one can look first at the subject matter or content of the image. In other words, what sense is made of the image? The content of the images can be read in ways that range from relatively low inference (like describing what

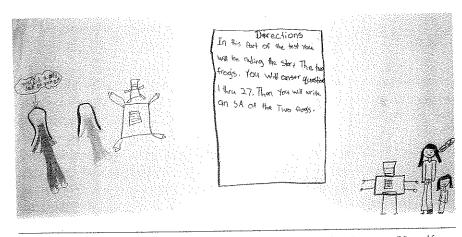**Table 10.1   Features of the Interpretation of Images**

	Readings of Images		Questions to Establish Credibility of Images
Focus on subject matter	• Literal reading	Low Inference ↓ High	• What are the physical features of the image? Who or what is portrayed? What is the setting?
	• Biographical reading		• What is the relationship of the image to current practices? To identities? How is the image socially situated?
	• Empathetic reading • Iconic reading		• What common experiences are invoked? • How does the image relate to bigger ideas, values, events, cultural constructions?
	• Psychological reading		• What are the intended states of mind and being?
Focus on image creation	• Technical reading • Editorial reading		• What are the design features? Are images scaled or mapped? • What values, knowledge are communicated by the author or creator?
	• Indexical reading • Spectatorship reading		• How does the image related to values of the time and place? • Where is the viewer situated in relation to the image?
Focus on audience/viewers	• Reading of effects • Reflexive reading		• What impact does the image have on viewers? • How do viewers see themselves within the interpretation? How does the image interact with biography?

is in the image) to much higher inference (such as imputing particular states of mind to people in images). The key overall concern is the degree to which the interpretation is supported by evidence from the images themselves. Table 10.1 lists four kinds of readings that may be part of the interpretation and accompanying questions to test the credibility of those interpretations. Take, for example, the common image of a map. The content of the map can be interpreted in a relatively straightforward, literal way—such as what place is portrayed, and what features of the place are included. But the map also lends itself to higher inference interpretations such as what kinds of places or things are valued, perhaps by being given prominence (museums, personal events), and what is devalued by its absence. Specifically, if one

looks at the common iconic map of the London Underground, the image portrays a sense of places in relation to one another, communicates how far apart places are, and evokes a sense of neatness and symmetry to the city. If one then compares this iconic map to the less common but geographically correct map of the London Underground, a different sense of place is communicated.

In my research about the meaning children make of the high-stakes testing experience in schools, we solicited drawings of kids in their classrooms. We found that when fourth-grade students were asked to draw pictures of their test-preparation activities, the urban students drew themselves with a teacher standing near them saying things like "good job" (see Figure 10.3), thus portraying a particular relationship between teacher and students in this context.

But images are not just of something, someone, or someplace. They are created, and in image-based research they can be created by the researcher/ evaluator, by the participants, or by a collaborative endeavor. Judging the credibility of the image and the study may also require attention to the creator. At one level, the credibility of the image can be judged by its technical qualities. If appropriate, does the image include elements of scale that provide adequate context for its interpretation? How is the image set up with regard to perspective? No image is a transparent window, and the creator's choices in its creation provide a point of view from which to understand the context. This point of view includes not only the technical decisions (black and white or color, line drawing or oil painting, distant or

Figure 10.3   Urban Fourth Grader's Drawing of Relationship Between Herself and Her Teacher

close up, front or back images) but also reflects the creator's cultural context and values in a particular time and place. In Table 10.1, these are referred to as editorial and indexical readings. The last sort of reading is one that considers the interpretation of the image in relation to the viewer: where does the image situate the viewer, the reader of the image?

Look again at Figure 10.2, Joseph's self portrait while taking the English Language Arts test. In this case, the creator of the image is known, which is not always the case, especially if images are found rather than created in the study context. Joseph can tell us things about what the image means, and he does in the interview excerpt included with his portrait. There are things about the drawing that Joseph may not be able to tell us, but that may be meaningful, such as the significance of the colors he chose, and the fact that Joseph looks directly at us out of his portrait.

Images imply a viewer, and image-based studies specifically define the viewer as the researcher/evaluator or the study consumer. Thus the credibility of the use and interpretations of images requires attention to audience. What impact do the images have on viewers? Are the researcher's interpretations coherent with the viewer's interpretations? How do the images and the interpretations of them interact with the biographies of consumers? Take Joseph's self portrait again as an example. We might ask, how do we see Joseph in the testing situation? and how do others understand Joseph's experience? I have a large matted version of this drawing, along with several others, on my office wall. They are pictures that draw people in, and inevitably visitors want to talk about them. Seeing these pictures, people ask me questions— about the kids, about the research, about why I used drawing. But they are also compelled to speculate about what they mean, to show empathy by feeling sad for kids like Joseph, to share in the satire inherent in some drawings, to be happy for the drawings that show pride in doing well on the test, and to tell stories of their own experiences with assessment in schools.

Credibility Is Dependent on the Extent to Which the Image-Based Study Contributes to New Knowledge

It is a truism to suggest that research and evaluation should result in new knowledge, new understandings of the focus of the study. In a traditional social science framework, a relatively small, particular, and focused audience is implied. Study results are reported in peer-reviewed journals and books that speak to others with similar interests to those of the study authors. New knowledge in this context is new to a group of colleagues working hard on inquiry about the same phenomena, and the study is indexed to an existing body of literature. This test of credibility may apply to image-based studies.

Attention is paid to whether the claims follow from the evidence, and whether the arguments are well-founded, defensible, and coherent.

The ubiquity and accessibility of images may suggest new contexts in which the credibility of the study to contribute to new knowledge might be judged. Because images have a populist character, image-based studies may be more accessible to study participants (and people like them), the public, and even policy makers, and therefore potentially broader contexts for knowledge generation may be important in considering the credibility of the study. Credibility in this more populist context might also involve questions about the image-based study's verisimilitude, authenticity, and usefulness as a catalyst for change. Instead of a rhetoric of scientific proceduralism, the rhetoric is one of believability, often a call to join the study author and participants on a journey to understanding and knowing (Wagner, 2004, p. 1500). In this context, the balance of interpretation to evidence will tip toward evidence, in contrast with the social science reports where interpretation is illustrated by evidence. This difference may contribute to the ambiguity of image-based studies, but also emphasizes the profoundly empirical nature of such inquiry.

This populist context of credibility does not seek legitimation within a body of literature created by a relatively small community of scholars, but rather seeks legitimation in its ability to communicate with diverse consumers. Evaluation studies are first and foremost about understanding the value and merit of something within a particular context, often with an eye to making things better. While image-based approaches are not a singular answer to how to do this well, it is the case that they may have certain advantages given this intent. Wagner (2004) suggests that inquiry should be multifaceted: "Being smart about that requires that we learn what we can from not only social scientists but also documentary photographers and filmmakers, at least some of whom celebrate both art and empiricism and aim for both telling images and telling truths" (p. 1505).

Conclusion: So What Counts as Credible Evidence?

It is too simplistic to assert that seeing is believing, but the fact that our eyes sometimes deceive does not obviate the credible knowing from doing and viewing image-based research and evaluation. This chapter has focused on how we establish the credibility of images and image-based research and evaluation, not because images are better or more compelling evidence but because this detailed examination illustrates that any evidence may be

credible and truthful. As suggested in the introduction to this chapter, the credibility of any evidence ought not to be established solely or even primarily by the methods by which it is generated. Whether one is doing ethnography or randomized experiments or survey research is not the key feature in establishing the credibility of the evidence generated. What counts as credible evidence is defined within inquiry traditions, within epistemological traditions. There are fundamental differences among these traditions, so the credibility of evidence or data within any particular tradition is first established within that particular tradition.

As we investigate and evaluate social phenomena, the larger context that eschews a hegemony of any particular tradition and any particular kind of evidence increases our potential knowledge of and methods of valuing those social phenomena. Images, correlations, effects sizes, vignettes—indeed, any and all evidence—must be taken seriously and examined in ways that reject the search for a single truth. Considering any and all evidence moves us to thinking about the usability and sensibility of evidence within the larger social context.

References

Becker, H. S. (1998). Visual sociology, documentary photography, and photojournalism: It's (almost) all a matter of context. In J. Prosser (Ed.), *Image-based research: A sourcebook for qualitative researchers*. London: Falmer.

Clark, C. D. (1999). The auto-driven interview: A photographic viewfinder into children's experiences. *Visual Sociology, 14*, 39–50.

Collier, J., Jr., & Collier, M. (1986). *Visual anthropology: Photography as a research method*. Albuquerque: University of New Mexico Press.

Epstein, I., Stevens, B., McKeever, P., & Baruchel, S. (2006). Photo elicitation interview (PEI): Using photos to elicit children's perspectives. *International Journal of Qualitative Methods, 5*(3), Article 1. Retrieved January 3, 2007, from http://www.ualberta.ca/~ijqm/backissues/5_3/pdf/epstein.pdf.

Feyerabend, P. (1975) *Against method*. New York: Humanities Press.

Freeman, M., & Mathison, S. (2005). *Urban and suburban students' reported experiences with high-stakes testing in New York State*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Greenfield, L. (n.d.). *Fast forward*. Retrieved January, 8, 2007, from http://zonezero.com/exposiciones/fotografos/lauren2/portada.html.

Harper, D. (2002). Talking about pictures: A case for photo elicitation. *Visual Studies, 17*, 13–26.

Hubbard, J. (n.d.). *Shooting back*. Retrieved February 6, 2007, from http://www.shootingback.org.

Kaplan, I., & Howes, A. (2004). Seeing through different eyes: Exploring the value of participative research using images in schools. *Cambridge Journal of Education, 34*(2), 143–155.

Loizos, P. (2000). Video, film, and photographs as research documents. In M. W. Bauer & G. Gaskell (Eds.), *Qualitative researching with text, image, and sound.* Thousand Oaks, CA: Sage Publications.

Mead, M., & Bateson, G. (1942). *Balinese character: A photographic analysis.* New York: New York Academy of Sciences.

Meijer, P. C., Zanting, A., & Verloop, N. (2002). How can student teachers elicit experienced teachers' practical knowledge? Tools, suggestions, and significance. *Journal of Teacher Education, 53*(5), 406–419.

Menzel, P. (1995). *Material world: A global family portrait.* San Francisco: Sierra Club Books.

Pullman, M., & Robson, S. (2007). A picture is worth a thousand words: Using photo-elicitation to solicit hotel guest feedback. *Cornell Hotel and Restaurant Administration Quarterly, 48*(2), 121–144.

Tufte, E. R. (1990). *Envisioning information.* Cheshire, CT: Graphics Press.

Tufte, E. R. (2003). *The cognitive style of PowerPoint.* Cheshire, CT: Graphics Press.

Tufte, E. R. (2006). *Beautiful evidence.* Cheshire, CT: Graphics Press.

Wagner, J. (2004). Constructing credible images: Documentary studies, social research, and visual studies. *American Behavioral Scientist, 47*(12), 1477–1506.

Wang, C. (1999). Photovoice: A participatory action research strategy applied to women's health. *Journal of Women's Health, 8*(2), 85–192.

Wang, C., & Burris, M. (1997). Photovoice: Concept, methodology, and use for participatory needs assessment. *Health Education and Behavior, 24*(3), 369–387.

Weber, S. (n.d.). *Concerning images.* Retrieved January 2, 2007, from http://iirc.mcgill.ca/

Werner, W. (2006). Reading pictures of people. In E. W. Ross (Ed.), *The social studies curriculum: Purposes, problems, and possibilities.* Albany: SUNY Press.

Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing: Merging auditory and visual worlds. *Neuron, 48*(3), 489–496.

11

Toward a Practical Theory of Evidence for Evaluation

Thomas A. Schwandt

Science is organized knowledge. Wisdom is organized life.

—Immanuel Kant

The field of evaluation remains captivated by the qualitative versus quantitative debate that engrossed it in the late 1970s and early 1980s. The latest manifestation of the debate is the controversy surrounding the nature and use of evidence in making decisions about the value (merit, worth, significance) of various evaluands. In this debate, defenders of experimental (read "quantitative") designs are squaring off against apologists for non-experimental (read "qualitative") designs in an argument about which is better capable of producing credible evidence. The original debate often generated more heat than light because it too frequently focused on quarrels about differences in methods rather than attending to more complex underlying issues at the intersection of the epistemology and social and political theory of evaluation. In much the same way, the present iteration of the debate is misguided because, once again, it centers on questions of method while ignoring more primary matters concerning what constitutes evidence and its responsible use.

To those engaged in the struggle to help create a better future,
including those whose efforts involve the use of applied research
and evaluation

What Counts as Credible Evidence in APPLIED RESEARCH and EVALUATION PRACTICE?

Stewart I. Donaldson
Claremont Graduate University

Christina A. Christie
Claremont Graduate University

Melvin M. Mark
The Pennsylvania State University

Editors

Contents

Preface

This volume addresses one of the most important and contentious i
challenging applied research and evaluation practice today: what c(
tutes credible evidence. We fear that history has been repeating itself of
with the earlier quantitative–qualitative paradigm war appearing aga
varied guises. These include public debates and serious professional ;
ments about federal research and evaluation design priorities, as well as $
ments from several professional associations regarding the role of randon
controlled trials (RCTs) and the existence of "gold standards" for evid
Much human capital and more than one friendship have been lost in h(
disagreements about what counts as credible and actionable evidence.
too, we fear, have been potentially productive conversations about the v;
ways in which applied research and evaluation can contribute to action
understandings that can assist in addressing important social needs.

One goal of this book was to address these thorny issues in a more ;
tive and productive manner. We hope that collectively the chapters in
book provide more light than heat regarding the fundamental chall(
of providing credible evidence, a challenge that faces contemporary ap[
researchers and evaluators in professional practice on a day-to-day basi

In an effort to accomplish this goal, a diverse and internationally renow
cast of applied researchers and evaluators were invited to participate in
Claremont Graduate University Stauffer Symposium focused on the ques
of "What Counts as Credible Evidence in Applied Research and Evalua
Practice?" This illuminating and action-packed day in Claremont was e>
rienced by more than 200 participants from a wide variety of backgrounc
evaluators, researchers, private consultants, students, faculty, and professio1
from many fields. The presenters all shared their latest thinking about hov
answer the question, and later each prepared a more detailed chapter to h
illuminate from their vantage point the key challenges about the nature
evidence in applied research and evaluation practice.