8

SEEING IS BELIEVING

Using Images as Evidence in Evaluation

Sandra Mathison

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

—Ad Rheinhardt, minimalist American painter

Images are evidence; this chapter explores the credibility and usefulness of images as evidence in evaluation. While there are unique features of images, their credibility exists within a larger framework for establishing facts and values. The credibility of evidence and the knowledge it generates is contingent on experience, perception, and social conventions. As such, knowledge changes 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 or data type. Embracing Feyerabend's (1975) anarchist epistemology, the notion that every idea or strategy however new or old or absurd is worthy of consideration, we can expand the possibilities for knowing the social world. Credible evidence is not the province of certain methods (such as observations or experiments) and cannot be expressed only one way (such as statistical averages or vignettes). To extend Feyerabend's notion, credible evidence should be

humanitarian, including 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 a commonsense realism prevails—we are reasonably certain at a given time of what is true, and we act on that knowledge with relatively high degrees of confidence. In some senses, certainty is based on cultural norms or intersubjective meanings, those things we collectively know even if that knowledge is tacit. But because knowledge is contingent, there are limits to its certainty, and what we know will inevitably change. The issue of certainty often 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.

Information becomes evidence within a context at both pragmatic and philosophical levels. Sometimes information becomes evidence because it is acquired in certain ways, such as through the use of randomized clinical trials (RCTs) within a post-positivist framework or through the telling of stories within a narrative framework. These are contexts of formal research and evaluation. But information becomes evidence through lived experiences, including professional practice. These are contexts of everyday life. Contextuality is important in two respects. First is the contextuality of the evidence itself, and the charge that something is taken out of context is always a serious challenge to 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 better able we are to judge its credibility. 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 rich data for understanding the social world and for forming and representing our judgments of that social world. Image-based research has a long history in cultural anthropology and sociology as well as the natural sciences, but it is nonetheless still relatively uncommon (Collier & Collier, 1986) and while used in program evaluation probably even less so than in research (Hurworth, 2004). Nonetheless, in this discussion, I want to explore both the uses of images and the possibilities for their use in evaluation.

I do not assert that images are better evidence than any other. Images should not be romanticized but neither should their value as data and knowledge be denigrated. Especially in Western industrialized cultures images are 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 are no more suspect than any other sort of data, such as numbers or text. Images, as 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.

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 seeing it. The 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 ball to move. Understanding DNA is another example. Most of us cannot see the genetic code, but we understand genetics when we look at siblings or parents and their children and note their similar appearance and behavior. Even when we do not directly see what we know, knowing is based on seeing and we typically call this *observation*. Seeing, even if in the form of indirect observation, 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 the central nervous system summarizes visual information in statistically optimal ways (Witten & Knudsen, 2005).

Situating Images as Evidence in Current Evaluation Approaches

Images are most likely to be used as evidence within participatory, empowerment, and responsive evaluation approaches. These are evaluation approaches that actively involve stakeholders in planning and implementing evaluation studies, as well as using evaluation results. Stakeholder involvement takes a particular form in this genre of evaluation—it must be broad and fundamental to decisions made in the evaluation process, often including defining evaluation questions, deciding what methods best answer those questions, deciding from whom to collect data, analyzing data, and disseminating and using evaluation results.

Participatory approaches to evaluation can be collaboration between an evaluator and program stakeholders, or program stakeholders might assume responsibility for the evaluation. In the collaborative scenario, evaluators and stakeholders work together to identify the evaluation focus, questions, and data collection, but the evaluator has responsibility for the execution and

quality control of the evaluation process. When program stakeholders assume responsibility, they call on evaluators to lend technical expertise if and as needed. The goal is increased stakeholder participation but also increased self-determination at individual, organizational, and/or community levels. Additionally, these approaches can be characterized by the following underlying motivations: (1) a pragmatic interest in making evaluation meaningful and useful, (2) a social constructivist epistemology, and/or (3) a political interest in social justice and fairness.

When evaluation perspectives, such as those described above, emphasize evidence meaningful to all stakeholders, and particularly service providers and recipients, there will likely be wide variation in evidence used in program evaluations. Careful attention is given to what is meaningful and sensible to not only evaluators but also to those who deliver and receive programmatic services. Encouraging program participants to use their own means to understand, analyze, and communicate the value of programs and services may include traditional social science research techniques, but will likely extend to more commonsense, pragmatic forms of evidence including, although not limited to, images. Potential benefits of images as evidence in evaluation are highlighting personal experiences, engaging emotions, surfacing assumptions, understanding complex and multi-dimensional experiences, and respecting cultural differences.

For example, photography is used for exploring and analyzing personal, cultural, and social experiences and the photographs, individually and collectively, can be evidence for judging the efficacy of interventions at each level (Wang, 1999). Sands, Reed, Harper and Shar (2009) used Wang's PhotoVoice to evaluate a school gardening program for fifth graders. Asking students, "If you were going to teach young children about growing food, what would you show them?" (p. 18) generated photographs used by the students for an analysis of core ideas about growing food, understandings that resulted from their participation in the program.

The California Endowment (2007) describes grassroots program evaluations that "give program participants a chance to tell their own stories in their own ways," which can take "many forms, from oral and written narratives to gesture, movement, art, music, movies and more" (p. 2).

The Center for Digital Storytelling has embarked on a project called Capturing Value directed at philanthropic grant makers wishing to expand the tools to convey the value of their work. Using a participatory evaluation framework, the Center invites work with funders wishing to use digital storytelling to explore what can be learned from funders' "investment in work for the public good" and that "can help others know how to move forward, sparking new ideas, inspiring new investments, guiding future choices" (Center for Digital Storytelling, n. d.). This initiative seeks to generate evidence that complements more traditional evaluation data by combining images and storytelling to engage audiences (other funders and policy makers, for example) through thought and feeling in ways that capture imagination, hold attention, and still communicate with nuance. Similarly, the Digital Clubhouse, with centers on both the U.S. east and west coasts uses the technology of digital storytelling to support people of all ages in considering how to improve the quality of their lives and communities. Both the Center for Digital Storytelling and the Digital Clubhouse provide instruction and support for using media technology, but the end goal is to use new media in service of giving authentic voice to individuals as they explore and try to enhance the social fabric of their lives, whether through a personal healing experience or a community development project.

Perhaps one of the most frequent contexts within which images are used in evaluation is when youth are primary stakeholders and a transformative participatory and empowerment evaluation approach is adopted. The school gardening project is an example of this, and the Digital Clubhouse emphasizes intergenerational involvement and several initiatives specifically

involve school-age children. Art Reach Toronto, a program that supports arts initiatives for engaging normally excluded youth, outlines a wide range of arts-based evidence and data-collection strategies to facilitate youth-focused program evaluations. In their Arts-Based Evaluation toolkit, creative strategies for expressing value are advocated to express complex ideas, to use multiple modes of expression, to build on creativity inherent in projects, and to generate data appropriate for communicating about value in reports (Art Reach Toronto, n. d.). PhotoVoice was the platform for youth involvement in the evaluation of Minnesota's Statewide Health Improvement Program. Youth from thirteen communities created a snapshot of Minnesota's health by focusing on how healthy eating affects them and their communities and what limits and supports physical activity and healthy eating in their community (Minnesota Department of Health, n. d.). These evaluation examples illustrate empowering youth to explore value with an eye to social improvement by using familiar and compelling technology tools, ways of exploring the world that may resonate more with their experience and skill than traditional social science tools.

What Is Image-Based Evidence in Evaluation?

Images in evaluation include found and evaluator- or participant-produced videos, photographs, drawings, collages, 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 a compelling representation of knowledge because they are fundamentally human ways of expressing experiences, values, and 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 evidence is any intrinsically visual data—such as photographs, video/film, drawings, cartoons, graphs and charts, typography, graphic art, graffiti, computer graphics, television, architecture, signs, clothing, 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 healthy communities reside in our mind, even if we have never seen 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.

Images as Data

In the first case, the images are themselves the data or evidence in an evaluation. The classic example of images as data comes from Mead and Bateson's (1942) photographic study of Balinese life. 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 an 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 8.1 shows a cartoon and Figure 8.2 a self-portrait by a fourth grader taking the state-mandated English Language Arts Test. The Performing Arts Workshop's quasi-experimental design evaluation of a school-based artist in residence program asked students in grades three to five to complete pre- and post-program surveys that included drawings as a response (Performing Arts Workshop, n. d.). These drawings, along with more traditional survey item data, allowed for an investigation of changes over time in student perspectives in and of art.

Figure 8.1   Cartoon Drawing as Data

Source: Freeman and Mathison (2005)

Figure 8.2   Joseph's Self-Portrait as Data

Joseph (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."

Source: Freeman and Mathison (2005)

Images to Elicit Other Data

Using images to elicit other data, such as 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). In general, evaluation participants are given the opportunity to take photographs related to the evaluation questions at hand. The specific subject matter is left to participants and the photographs become the physical manifestation of what participants pay attention to and, by inference, 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. In assessing customer satisfaction with a hotel, Pullman and

Robson (2007) asked guests to photograph things 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. Evaluators who bring a theoretical framework to their work may take photographs or videos that focus the elicitation on specific evaluation issues. White, Sasser, Bogran, and Morgan's (2009) use of Photolanguage in evaluation illustrates this strategy. Photolanguage "uses black-and-white photographic images specifically chosen for their aesthetic qualities; their ability to stimulate emotions, memory, and imagination; and their capacity to stimulate reflection in the viewer. Evaluators may find Photolanguage provides a valuable tool for use with evaluation participants who experience barriers to involvement, actual or perceived." Similarly, in a study of a therapeutic camp for children with cancer, investigators hypothesized that the nature of physical space was a critical dimension of the camp's therapeutic quality. Researchers took photos of camp spaces and used them as elicitation devices in interviews with campgoers to explore this idea (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).

Images as Representation

The third use of images is to represent knowledge generated within an evaluation. Visual images can illustrate the interpretive text—a legitimate and particular use of images. This use of images in 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. The idea of using a broad

range of communicative devices is not new in evaluation (Torres, Preskill, & Piontek, 2005), and images can augment or be the primary communicative device for reporting findings.

Media, other than text, have become increasingly accessible and communicable with the internet, such as Kaplan and Howes's (2004) web-based compilation of photographs taken by secondary school students and staff. The Kellogg Foundation summarized the issues from evaluations of rural community leadership programs (what the Foundation calls *cluster evaluation*) by developing a short video combining video images of rural communities, interview clips with program participants, and voice-over narration (Kellogg Foundation, 2013).

While not specific to evaluation, the Kids With Cameras projects illustrate the use of photographs as representation, made popular by the movie *Born into Brothels.* While this project uses photography as an empowering skill for these children, their photographs also stand as representations of their life, a representation that needs no further interpretation or embellishment. The idea is that participant photography permits participants to speak directly and empowers them 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 Lauren Greenfield's (n. d.) representation of growing up in Los Angeles in *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) illustrate images as knowledge. While examples of photo essays in evaluation reporting are scarce, the Asian Development Bank (ABD) uses this strategy extensively, including for reporting evaluation findings. The evaluation of ABD social protection programs includes a more traditional written report augmented by a photo essay (Asian Development Bank, 2012).

While it is easy to focus on the numbers represented in graphs, tables, and charts, these are also images, ones used by all evaluators, and there are many resources providing guidance for and examples of their effective use in evaluation reporting. Innovative ways of summarizing data and information into images are emerging with the increased use of infographics and digital dashboards in evaluation as well as attention to presentation aesthetics. Big Brothers Big Sisters used infographics to summarize Youth Outcomes Survey findings, an image that illustrates program goals, criteria for success, and outcomes (see Figure 8.3). Infographics go beyond summarizing data in a table or chart and employ graphic design to facilitate visualization of data, often communicating large data sets in simple, intuitive ways—Tufte's work is significant and instructive 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. He describes this as envisioning information, which is at "the intersection of image, word, number, art" (Tufte, 1990, p. 9). The potential of dynamic infographics for evaluation reporting are illustrated in the amazing video of Hans Rosling's *200 Countries, 200 years, 4 Minutes,* an animation of huge amounts of global health data in real space. (See Evergreen Data for a wide range of data representation ideas in evaluation; http://stephanieevergreen.com/)
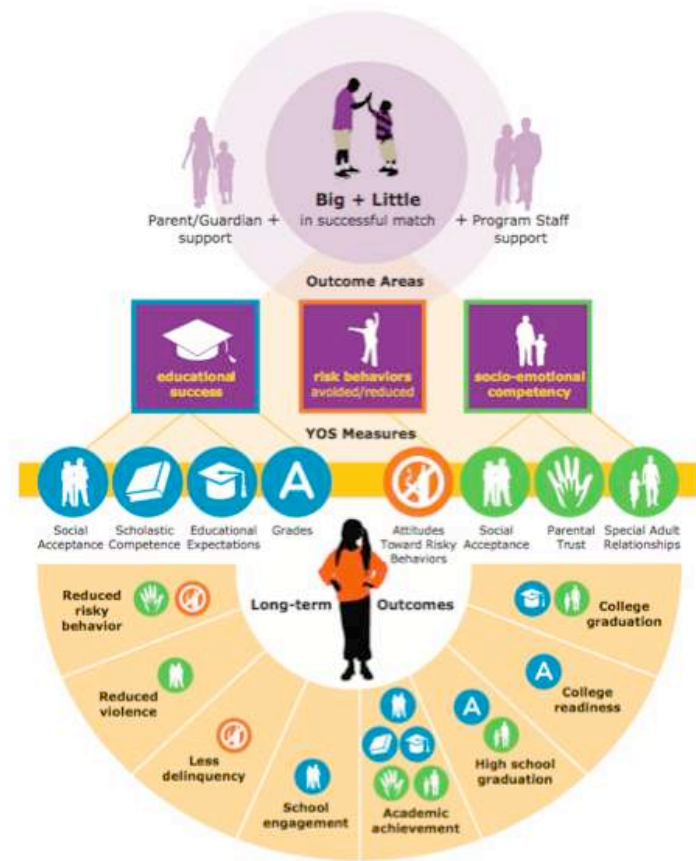
Figure 8.3 Infographic of Evolution Criteria and Program Outcomes for Big Brothers Big Sisters Student Outcome Survey

Source: Wheeler and Valentino (2013)

Credibility of Images in Evaluation

All data are susceptible to manipulation, distortion, or misuse, but images no more so than any other data. Just as statistics can distort meaning, so too can images distort meaning. *Photoshopped* has become a verb indicating an image has been digitally changed to deliberately distort meaning, but its meaning extends to other uses, such as when the media focuses on celebrity breakups rather than Hurricane Katrina victims or when history textbooks exclude

embarrassing moments of racism or sexism. The 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 image manipulation for whimsical, commercial, and political purposes, and the credibility of images as evidence 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 evaluators are no more likely to manipulate image data than they are to manipulate statistical data.

Four considerations for establishing the credibility of images are discussed below: (1) quality of the evaluation design, (2) attention to context, (3) adequacy of the image from multiple perspectives, and (4) the contribution images make to new and useful knowledge.

Evaluation Design Provides the Framework for Credibility

In part, the credibility of any data can be judged by reference to the evaluation design (Wagner, 2004). There is a debate about what constitutes the best evaluation design, marked by a recent advocacy for RCTs. 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 design, the credibility of images can be established within the context of whatever evaluation design is used.

Image-based evaluation can operate within a post-positivist framework where image data are treated as any other kind of data. The credibility of the evidence, and the threats to credibility, are established procedurally. Evaluation studies intend to answer questions or test hypotheses. Whether the evidence collected provides credible answers 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 and what they represent), the sampling of images from a knowable population in an explicit and defensible way, and whether the analysis follows understood conventions.

A participatory- or empowerment-oriented 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, or program 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 evaluator and participants) might be the focus.

Regardless of the underlying epistemologies, the credibility of images in evaluation can be judged by the defensibility of the evaluation design. Truth or credibility is, at least in part, established through conventions. While different camps of inquiry accept different conventions for establishing credibility, there is a set of conventions to which a particular image-based evaluation study can be compared. Image-based evaluation, as with any evaluation, 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 because the photographs sometimes 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 for understanding 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, p. 74) **[AU: Page number for this quote?]**

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, as do 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 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 images may also imply 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, detailed images.

Adequacy of the Interpretation of the Image

The adequacy of the interpretation of images is the third aspect of credibility and can be understood in terms of the focus on the subject matter, the creation, or the viewers of the images (Werner, 2006). One or the other of these three foci may be more relevant in establishing the credibility in any particular evaluation study, but at the heart of the adequacy of the interpretation is a sense of authenticity. Table 8.1 summarizes each foci, including key questions one might ask to establish the credibility of an image or an image-based study.

Table 8.1 Features of the Interpretation of Images

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

In assessing the adequacy of the interpretations of images, one can look first at the content of the image. In other words, what sense is made of the image? The content can be read in ways that range from relatively low inference (such as describing what is in the image) or much higher inference (such as imputing particular states of mind to people in images). The key concern is the degree to which the interpretation is supported by the images themselves. Table

8.1 lists the 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 map's content can be interpreted in a relatively straightforward, literal way—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. 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 distances between places, 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.

But images are not just of something, someone, or someplace. They are created by the evaluator, by other stakeholders, or are a collaborative endeavor. Judging the image's credibility 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 for understanding the context. This point of view includes not only the technical decisions (black and white or color, line drawing or oil painting, distant or close up, front or back images) but also reflects the creator's cultural context and values in a particular time and place. In Table 8.1, 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 8.2, Joseph's self-portrait while taking the English Language Arts Test. In this case, the image creator is known, which is not always the case especially if images are found rather than created. Joseph can tell us things about what the image means, and he does so in the interview excerpt. 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, thus the interpretations of images require attention to audience. What impact do the images have on viewers? Are the evaluator and stakeholders' interpretations coherent with the viewer's interpretations? How do the images and the interpretations of them interact with the biographies of viewers? 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, in my office. These pictures draw people in, and inevitably visitors 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 speculate about what they mean, show empathy by feeling sad for kids like Joseph, share in the satire inherent in some drawings, share the happiness in doing well on the test, and tell stories of their own experiences with assessment in schools.

Credibility Is Dependent on the Extent to Which Image Contributes to Useful Knowledge

It is a truism that evaluation should result in new knowledge, new understandings about the evaluand. In a traditional social science framework, a relatively small, particular, and focused audience is often implied. Study results are reported in peer-reviewed journals and books read by others with similar interests. 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.

Evaluation assumes a broader audience though, including policy makers, funders, communities, practitioners, and service recipients. Evaluation is complex, and the field has moved well beyond the simplistic instrumental-use notion to more nuanced and varied uses of evaluation and evaluation findings (Kirkhart, 2000; Preskill & Caracelli, 1997; Weiss, 1979). In some arenas, policy makers have regrettably and effectively redefined credible and useful evidence by specifying which evaluation designs are best, resulting in what Weiss, Murphy-Graham, Petrosino, and Gandhi (2008) call *imposed use*. Results of evaluations that use experimental designs are categorically defined as more useful, but at the end of the day imposed use still begs questions about the credibility of evidence, discounts local knowledge of program effects, and privileges bureaucratic procedures over arguments to demonstrate usefulness (Weiss et al., 2008).

The ubiquity and accessibility of images may suggest alternatives for judging the credibility and usefulness of evaluations that rely on images as evidence. Because images have a populist character, image-based evaluations and evaluation findings may be more accessible to study participants (and people similar to them), the public, and even policy makers, and therefore, potentially broader contexts for knowledge generation may be important in considering a study's credibility. Credibility in this more populist context might also involve questions about verisimilitude, authenticity, and usefulness as a catalyst for change. Instead of a scientific proceduralism rhetoric, the rhetoric is one of believability, often a call to join the evaluator and other stakeholders 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 populist context of credibility does not seek legitimation within a body of literature created by a relatively small community of scholars and does not fit with the imposed use now common in the federal government arena but rather seeks legitimation in its ability to communicate with diverse consumers. Evaluation studies are first and foremost about understanding value and merit within a particular context, often with an eye to making things better. While images are not a singular answer to how to do this well, 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: Do Images Count as Credible and Useful Evidence?

It is simplistic to assert seeing is believing, but because our eyes sometimes deceive us does not obviate the credible knowing from using and viewing images in evaluation. This chapter focused on how we use and establish the credibility of images in program evaluation not because images are better or more compelling evidence but because any evidence may be credible, truthful, and useful. As suggested in the introduction, the credibility of 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 only feature in establishing the credibility of the evidence generated. What counts as credible evidence is defined within inquiry traditions, within epistemological traditions. In evaluation theory and practice,

fundamental epistemological differences exist, so the credibility of evidence within any particular tradition is first established within a particular tradition.

As we investigate and evaluate social phenomena, we should eschew the hegemony of any particular tradition and assume all kinds of evidence might increase our knowledge of and methods of valuing 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 conclusion about what is good and right. Considering any and all evidence fosters thinking about the usability and sensibility of evidence within the larger social context.

References

Art Reach Toronto. (n. d.). *Arts-based evaluation 101.* Retrieved April 16, 2014 from

http://www.artreachtoronto.ca/toolkits/the-goal-toolkits/art-based-evaluation/

Asian Development Bank. (2012). *Special evaluation study on ABD's social protection strategy.*

Retrieved April 9, 2014, from http://www.adb.org/sites/default/files/ses-

socialprotection.pdf

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, England: Falmer.

California Endowment. (2007). *Storytelling approaches to program evaluation: An introduction.*

Los Angeles, CA: Author.

Center for Digital Storytelling. (n. d.) *Capturing value.* Retrieved on April 16, 2014 from

http://www.storycenter.org/capturing-value/

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, NM: 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 April 9, 2014, from http://www.ualberta.ca/~iiqm/backissues/5_3/PDF/epstein.pdf

Feyerabend, P. (1975). *Against method.* New York, NY: 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 in Montreal, Quebec, Canada.

Greenfield, L. (n. d.). *Fast forward.* Retrieved April 16, 2014, 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 April 16, 2014, from http://www.shootingback.net

Hurworth, R. (2004). The use of the visual medium for program evaluation. In C. Pole (Ed.), *Seeing is believing? Approaches to visual research* (pp. 163–180). Oxford, England: Elsevier Ltd.

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.

Kellogg Foundation. (2013, February 5). *God made a rural*. Video retrieved from

      http://www.youtube.com/watch?v=0LusMW84c5Y

Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New*

      Directions for Evaluation, 88, 5–23.

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, NY:

      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, CA: Sierra Club

      Books.

Minnesota Department of Health. (n. d.). *A snapshot of health in Minnesota: Youth PhotoVoice*

      project. Retrieved on April 15, 2014 from

      http://www.health.state.mn.us/healthreform/ship/events/policyconference/photovoice.pdf

Performing Arts Workshop. (n. d.). *ARISE student survey*. Retrieved on July 9, 2013 from

      http://www.performingartsworkshop.org/pages/resources_administrators.html

Preskill, H., & Caracelli, V. (1997). Current and developing conceptions of use: Evaluation use

      TIG survey results. *Evaluation Practice, 18,* 209–225.

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.

Sands, C., Reed, L., Harper, K., & Shar, M. (2009). A PhotoVoice participatory evaluation of a school gardening program through the eyes of fifth graders. *Practicing Anthropology, 31*(4), 15–20.

Torres, R. T., Preskill, H., & Piontek, M. E. (2005). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations.* Thousand Oaks, CA: Sage Publications.

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 April 16, 2014, from http://iirc.mcgill.ca/

Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review, 39,* 426–431.

Weiss, C. H., Murphy-Graham, E., Petrosino, A., & Gandhi, A. G. (2008). The fairy

godmother—and her warts: Making the dream of evidence-based policy come true.

American Journal of Evaluation, 29(1), 29–47.

Werner, W. (2006). Reading pictures of people. In E. W. Ross (Ed.), *The social studies

curriculum: Purposes, problems, and possibilities.* Albany, NY: SUNY Press.

Wheeler, M., & Valentino, S. (2013, April 23). *Using infographics to translate evaluation

evidence: The 2013 BBBS youth outcomes report.* Retrieved on August 1, 2013, from

http://chronicle.umbmentoring.org/research-corner-using-infographics-to-translate-

evaluation-evidence-the-2013-big-brothers-big-sisters-youth-outcomes-report/

White, R., Sasser, D., Bogran, R., & Morgan, J. (2009). Photos can inspire a thousand words:

Photolanguage as a qualitative evaluation method. *Journal of Extension, 47*(3). Retrieved

April 15, 2014, from http://www.joe.org/joe/2009june/iw1.php

Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing: Merging auditory and visual

worlds. *Neuron, 48*(3), 489–496.