Toward Evaluating Art Museum Education at the Art Gallery of Ontario

Agnieszka Chalas, Queen's University
a.chalas@queensu.ca

Abstract Over the past three decades, the museum education field has seen a rise in the frequency of program evaluation. In this paper, I convey little known information about program evaluation at the Art Gallery of Ontario by presenting my findings from an interview I conducted with Judy Koke, the gallery's Chief of Public Programming and Learning. Our discussion highlights both the barriers the AGO has faced on their journey toward evaluating programmatic value and the strategies the gallery has employed in an effort to enhance its internal evaluation efforts. A brief overview of program evaluation in museums provides the background to this discussion.

Keywords: Éducation musée d'art; évaluation du programme.

"If museums are to survive long term, they need to be able to articulate why they matter in our communities…and I think [visitor research and evaluation] is the path to getting there."—Judy Koke

Introduction

Evaluation[i] is critical to determining what learning occurs as a result of program participation. As a long-time art museum educator and former Director of Public Programs and Education at the Kitchener-Waterloo Art Gallery, I am well aware that program evaluation is still a more unusual occurrence than it ought to be, especially in small to mid-sized institutions with limited capacity for evaluation. This observation is also supported by conversations I have had over the years with a variety of museum professionals working in regional public art galleries and museums across Ontario. As one of my colleagues recently noted, "You'll be hard-pressed to find a smaller art gallery with either the staffing or the funds to dedicate solely to program evaluation" (art museum director, personal conversation, March 3, 2015).

I knew that, despite this, art museums[ii] in large urban areas would have more resources to draw upon. Being one of North America's largest art institutions, I was curious to know: How active is the Art Gallery of Ontario (AGO) in evaluating their educational program offerings? This starting point was born of my own evaluation struggles, a social constructivist epistemology, and a pragmatic interest in empowering art museum educators with the skills necessary to both *conduct* their own evaluations and make *use* of evaluation findings and processes. To answer my question, I interviewed Judy Koke, Chief of Public Programming and Learning at the AGO. Having a broad base of experience with education and evaluation in museums, Judy is credited with "developing and implementing a pan-institutional program of visitor research and evaluation and leading a cultural change to support data-driven, visitor-centred decision making" (Kansas City InfoZine, 2011, para. 3). She was therefore in an excellent position to speak to the role of program evaluation in the AGO's overall visitor research[iii] agenda.

I selected a semi-structured, open-ended phenomenological interview approach (Patton, 2015) to design a detailed interview guide consisting of ten open-ended items. Lasting roughly 45 minutes, the interview was conducted on March 13, 2015 via Skype. The dialogue was recorded using a digital recorder and later transcribed in full. A standard inductive analysis (Patton, 2015) allowed for the emergence of the following three themes, which serve as an interpretive framework for the interview findings: (a) Embracing Evaluative Thinking: A Strategy for Institutional Change, (b) On Evaluation: Barriers and Breakthroughs, and (c) Picturing the Future of Evaluation: The AGO and Beyond. The interview data is presented in a conventional format that intersperses verbatim quotes in summaries in an effort to assist readers in keeping track of themes (Thody, 2006). Before discussing each theme in turn, I present the

context for the interview by way of a brief introduction to program evaluation theory and practice within the art museum context.

Educational Program Evaluation in Art Museums: The Landscape, the Challenge, and the

Reward

Encompassing a wide variety of approaches, program evaluation is a professional practice dedicated to making judgements about programs. For the purpose of this paper the following definition as presented by Patton (2008) is adopted:

> Program evaluation is the systematic collection of information about the activities, characteristics, and results of programs to make judgments about the program, improve or further develop program effectiveness, inform decisions about future programming, and/or increase understanding. (p. 39)

It is important to note that evaluation should not be confused with other forms of inquiry or reduced to one approach or purpose. While the aim of research is to generate new knowledge, evaluation's ultimate purpose is to inform decisions (e.g. about processes, resources, strategies, or improvements). Since its emergence in the mid-1960s (see Hein, 1998), the field of program evaluation has unleashed a profusion of approaches and seen qualitative and mixed-methods approaches, in particular, rise in stature (Chouinard, 2013). Today, the program evaluation field is characterized by significant diversity in both approaches and theoretical perspectives (Alkin, 2004; Shadish, Cook, & Leviton, 1991). There exist, throughout the program evaluation literature, a number of classification systems that compare approaches and serve to guide practitioners when choosing one to use. The evaluation theory tree, presented in Alkin's book *Evaluation Roots* (2004; 2013), is one such system. This conceptual framework categorizes evaluation theorists based on the emphasis they place on one of three dimensions or 'branches'

(*use*, *methods*, and *valuing*) with the understanding that all evaluation theorists are concerned with methods and recognize that evaluation involves judgment and the use of findings to affect programs (Christie & Alkin, 2008). Chouinard (2013) offers a more binary classification system, distinguishing between what she terms 'technocratic' approaches to evaluation grounded in quantitative (positivist/post-positivist) methodologies and participatory approaches grounded in qualitative (constructivist/interpretive) ones (e.g., Cousins & Chouinard, 2012; Fetterman & Wandersman, 2005; Rodríguez-Campos, 2005). The former is characterized as an accountability-driven approach and the latter by its engagement of non-evaluator stakeholders in the evaluation process (see also MacDonald's, 1974 and Schwandts', 2009 classification systems).

While some might argue that evaluation approaches based on a constructivist paradigm are more compatible with the learner-centred theories that undergird art museum education, Alkin (2004) rightly points out that "there are, in fact, multiple acceptable ways to pursue evaluation" (Editor's Introduction, para. 7). Hein (1998) further reminds us that "all professionals in the field recognize the need for using multiple methods, and especially, for using the appropriate method for the research questions of interest" (p. 134).

Over the past three decades, the museum education field has seen a rise in the frequency of program evaluation. Museum practitioners' desires to be more visitor-centred and funders' emphases on accountability have fuelled the demand for evaluations of publicly funded programs (Gorman, 2007; Hein, 1998; Korn, 2007). Today, museums of all varieties, including art museums, conduct program evaluations in whatever capacity their resources allow, with many museums either hiring in-house evaluators or external consultants to provide evaluation services (Falk, 2000; Luke & Ancelet, 2014). And, while many museums operate on a *contractual* evaluation model, wherein control of technical decision-making surrounding the evaluation

design and data analysis rests with the evaluator, the use of participatory evaluation approaches is on the ascent, marking an important trend in the field of evaluation (Rodríguez-Campos, 2012). Such approaches see museum staff work alongside professional evaluators in an ongoing and bottom-up way, enhancing their capacity to conduct subsequent evaluations independently. King (2005) elaborates on this development, suggesting, "by taking part in studies (i.e., learning evaluation by doing it) participants may better understand the evaluation process and gain program evaluation skills" (p. 292). Defined in this way, program evaluation is conceptualized as an empowering process and a pedagogical undertaking in itself (Cousins & Earl, 1992; Schwandt, 2003). The John G. Shedd Aquarium in Chicago (Kubarek, 2015) and the Science Museum of Minnesota in St. Paul (Grack Nelson & Callahan Schreiber, 2009) are two examples of museums that have recently employed a participatory approach when evaluating their educational programs. Although, these program evaluations were conducted in science rather than art settings, the growing trend in the field towards participatory evaluation is reflective of museum practice in other contexts. Still other museums interested in evaluating their own programs have engaged more directly in institution-wide Evaluation Capacity Building (ECB)—a practice, related to but distinct from program evaluation, aimed at building evaluation culture and praxis in organizations (Compton, Baizerman, & Stockdill, 2002; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012). ECB activities are planned, both ongoing and integrated into an organization's operations, and can occur either within or outside of actual evaluation projects (Compton, et al., 2002; Preskill & Torres, 1999; Rowe & Jacobs, 1998). Further, these activities may be delivered in a myriad of different ways from face-to-face training sessions to webinars, and can be directed at the individual or organizational level (Labin et al., 2012).

The evaluation and museum literature provides a number of frameworks for both measuring museum learning and building internal evaluation capacity (e.g., Falk & Dierking, 1992, 2000; Gorman, 2007; Hooper-Greenhill, 2004; King & Volkov, 2005; Kirkpatrick, 2003). Despite the availability of an increasing number of such frameworks, responding to accountability demands for program evaluations has nonetheless posed some significant challenges for art museums (Kelly, 2004). For example, while larger institutions can afford to hire professional evaluators, the task of evaluation in small to mid-sized museums falls to what King and Volkov (2005) term *accidental evaluators*—program staff with little or no training in evaluation who are nonetheless required to demonstrate measurable outcomes of programs. Beleaguered by other responsibilities and challenged with a lack of finances, many such practitioners would prefer to devote their limited time and resources to program development and delivery rather than evaluation (King & Volkov, 2005). Many more are uncomfortable with the language of evaluation, erroneously associating *all* program evaluation approaches with positivist science despite the rich and varied methodological options available (Hooper-Greenhill, 2004). In today's climate, accountability is not optional. Accordingly, as we move into the future of art museum education, it is unlikely that the demand for accountability-driven evaluation will be tempered. Increasing art museum educators' understandings of program evaluation through direct or indirect (e.g., participatory evaluation approaches) ECB efforts is therefore of paramount importance to such museums.

Ongoing debates in the field over what to evaluate and which measures to use pose an additional challenge (Dierking, Ellenbogen, & Falk, 2004; Gruber & Hobbs, 2002). Too often, when evaluation does take place, outputs like attendance figures, repeat visitation, and visitor satisfaction rates are some of the most predominant measures museums use to evaluate the

success of their programs. However, as Korn (2008) observes, the value of museum programs extends far beyond such outputs, which do little to describe whether or not programs meet their learning-related outcomes or how participants benefit. Current scholarship argues that, if museums are to demonstrate how art museum education impacts on program participants and facilitates learning, they will need to focus on performance indicators other than outputs (Falk & Sheppard, 2006; Luke & Ancelet, 2014; Preskill, 2011; Suchy, 2004; Worts, 2006). This, Adams (2012) argues, will require museums to think more broadly and creatively about the evaluation methods they employ and the type of data they collect.

Notwithstanding these challenges, which are by no means exhaustive, interest in program evaluation within the field of art museum education is growing. Art museum educators are slowly embracing program evaluation's potential to both generate accountability evidence and improve educational programming. In addition to honing important evaluation skills, numerous authors further suggest that program evaluation can also play a key role in facilitating ongoing organizational learning (Hein, 1998; Korn, 2007; Senge, 1999; York, 2005). By combining program development with evaluation, they argue, the museum engages in an iterative process of learning about its 'organizational self'. Moreover, evaluation can increase the status of educational departments by promoting their work within the hierarchy of the museum and helping to attract support for new and expanded programming (Gorman, 2007; Korn, 2007; Linn, 2000). A number of scholars have recently expressed concern over the sustainability of the museum sector. Jacobsen (2010) perhaps expresses this concern most strongly when he says that the museum "is in danger" (p. 281). If art museums are to be understood and promoted as integral parts of a society-wide learning infrastructure, they will need to demonstrate effectiveness in their educational programming. By measuring learning (broadly defined),

regardless of the approach used, program evaluation can enable art museums to better position themselves in terms of their educational value to the public.

Embracing Evaluative Thinking: A Strategy for Institutional Change

The AGO has transformed itself in significant ways over the last several years and not just in terms of its much-publicized 2008 Frank Gehry architectural re-design. Judy Koke attributed the shift in the institution's culture towards embracing 'evidence-informed visitor-centred decision-making' as an important tool in this transformation. Despite being instigated by the museum's leadership, this shift did not occur overnight. "Initially there was no pan-institutional conversation about what we were learning from [evaluation]," Judy noted, reinforcing Luke and Ancelet's (2014) observation that institutional change takes time. She went on to describe how she engaged a full range of museum staff in adopting an evaluative mindset:

What I think really made the biggest difference was that I put together a Visitor Research Committee. And Marketing was on it. And Membership was on it. And Visitor Services was on it. And Education was on it. We all started to talk about our data across the whole institution and together defined the big institutional questions we're trying to answer… so it's evolved to become a much less 'siloed' process and much more of a shared process.

Despite this, Judy lamented that the curatorial staff still hold predominantly negative perceptions about evaluation, confirming Preskill's (2011) suggestion that "evaluation is often seen by others as threatening, anxiety producing, and potentially punishing" (p. 97). Judy spoke to this, explaining:

The exhibitions people, the project managers, the designers, the educators on a project, the programmers…all are very interested in evaluation and find it a useful tool, but there's not that same appreciation in the curatorial division. I think sometimes the curators in an art

museum feel like we're evaluating their idea, which makes them defensive, understandably. I'm hoping that that will continue to change.

As much as ongoing work needs to be done to help curators and educators in art museums see how evaluation can inform their exhibition and educational program development process in a positive way, evaluation is much more successful when it is seen as a team effort. Adams (2012) argues that, "getting education, curatorial, visitor services, marketing, exhibition design, and membership to work together, while often challenging, yields amazing results" (p. 30).

Having worked in a variety of museum contexts, including natural history and science museums, Judy believes that program evaluation in art museums lags behind other institutions—a view that is echoed by others including Luke and Knutson (2010) who assert that more work is needed to articulate the range of effects and outcomes attributable to art museum education. Art museums might, therefore, learn from other museums that have embraced evaluative thinking and successfully mainstreamed the evaluation process into their day-to-day operations.

On Evaluation: Barriers and Breakthroughs

From my conversation with Judy, it was clear that the AGO is quite active in evaluating its exhibition projects, but considerably less so when it comes to evaluating its educational programs. "We're trying to measure learning very broadly defined and we're not very strong yet at evaluating our public programs…it's not like we don't do it. We just don't do it as well as we should," she admitted. That said, Judy recognized the value and utility of educational program evaluation, and went as far as to say, "If museums are to survive long-term, they need to be able to articulate why they matter in our communities…and I think [visitor research and evaluation] is the path to getting there." Judy identified a number of barriers to implementation with a dearth of funding being the greatest contributing factor to program evaluation's slow uptake in Canadian

art museums. She explained that, "The Institute for Museum and Library Services [in the US], for example, requires 10% of any grant proposal to be dedicated to evaluation, so it's really healthily funded in the States. It's not funded here. When you spend money on evaluation it's seen as taking money away from program delivery." Although the funding landscape is beginning to show evidence of increased resources for evaluation (e.g., the Ontario Trillium Foundation), Canadian funders, as a whole, could play a greater role in supporting program evaluation by requiring grantees to direct a portion of the total grant value towards the comprehensive evaluation of programmatic impacts.

Other barriers mentioned included a lack of time and personnel in addition to the marginalization of educational endeavours—the latter challenge being, in Judy's opinion, unique to the art museum context. "If you ask most people what our biggest product is, they will say 'the exhibitions'. So, it's not surprising that maybe that's where the resources are first allocated—towards evaluating the success of our temporary ticketed exhibitions," Judy clarified. These barriers reinforce a recent observation made by Adams (2012) that, "While the 'want to' in the practitioner's desire to conduct evaluation is much stronger than ever before, the 'can do' is a different matter" (p.28).

Museum-university collaborations, Judy suggested, have the potential to move this desire into action and enhance museums' evaluation efforts. "We have a partnership with the *Master of Museum Studies* program at the University of Toronto. One of the courses they teach is an evaluation course and we co-teach it with them and [students taking the course] do evaluation projects [at the gallery]," she said excitedly. Owen and Visscher (2015) claim that there is great value in this approach. They write, "Conducting audience research through university and museum partnerships has the potential to be mutually beneficial—providing museums with

valuable insights into the audiences they serve and giving students hands-on experience in a professional setting" (p. 71). Thus, for art museums interested in integrating the evaluation process into their ongoing work, museum-university collaborations can be a viable starting point.

Picturing the Future of Evaluation: The AGO and Beyond

When asked about her future research agenda, Judy identified placing a greater emphasis on educational program evaluation as a future goal for the gallery, which she felt optimistic about attaining. "Absolutely doing more and better program evaluation is part of it," Judy told me. She added, "It's in the realm of possibility…there's a real appetite for making data-driven decisions here now so there's no shortage of desire for program evaluation. It's just that we haven't quite caught up in terms of how we resource it." Judy also expressed excitement about building museum networks as a strategy for increasing museums' capacities to conduct evaluations. While pointing to the Denver-area Evaluation Network (DEN, 2015) as an example of a group of museums successfully working together to collect data as a network, she said:

> I think there's real power in collaboration! What if four or five museums got together and decided what are two or three big evaluation questions that they would love to have answered and then work together to decide how to explore those questions? I feel like we could do that! [Collaboration] makes [evaluation] seem more doable.

As others have noted (Adams, 2012; Steele-Inama, 2015), networks of museums are increasingly working collaboratively around evaluation—learning how to conduct their own evaluations as well as carrying out pan-institutional evaluations together. Steele-Inama (2015) believes that forming evaluation networks for museum professionals can "positively influence evaluative thinking, implementation, and use" (p. 80). Museum networks then, such as the Denver-area Evaluation Network, are a promising way to increase evaluation capacity and can

serve as models for both the AGO and other art museums across the country.

Conclusion

As Falk and Dierking (2000) rightly point out, "that people learn in museums is easy to state, harder to prove" (p. 149). An examination of the AGO's efforts to build evaluation into its operations, highlights some of the barriers that art museums face on their journeys toward evaluating programmatic impact. Although the AGO is not evaluating their educational programs to the degree that I had expected, the headway made to date, especially with respect to institutionalizing the gallery's evaluation function, marks a step in the right direction and provides a springboard for future success in this important area of practice. Numerous authors have called for art museums to start evaluating their educational programs, and to evaluate them well (Falk, 2000; Hein, 1998, 2005; Parsons, 2004; Sheppard, 2000, 2010; Worts, 2006). The AGO's future efforts, therefore, should demonstrate a commitment to a more comprehensive approach to evaluation—one that examines the outcomes of both exhibitions and educational programs. In doing so, the AGO will be able to show a much broader picture of the gallery's institutional value than can be reported by evaluating the impact of its exhibitions alone. And, in future, when asked the question: "What kind of learning occurs as a result of program participation?" the AGO will be better positioned to provide evidence-informed answers. It was a pleasure to interview Judy Koke on this subject and I am excited to see how program evaluation at the AGO evolves in the coming years.

References

Adams, M. (2012). Museum evaluation: Where have we been? What has changed? And where do we need to go next? *Journal of Museum Education, 37*(2), 25–36.

Alkin, M. C. (2004). *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage.

Alkin, M. C. (2013). *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed.). Thousand Oaks, CA: Sage.

Christie, C. A., & Alkin, M. C. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation*, *34*, 131–135.

Chouinard, J. A. (2013). The case for participatory evaluation in an era of accountability. *American Journal of Evaluation*, *34*(2), 237–253. doi: 10.1177/1098214013478142

Compton, D., Baizerman, M., & Stockdill, S. H. (Eds.). (2002). *New directions for evaluation: No. 93. The art, craft and science of evaluation capacity building*. San Francisco, CA: Jossey-Bass.

Cousins, J. B., & Chouinard, J. A. (2012). *Participatory evaluation up close: An integration of research-based knowledge*. Charlotte, NC: Information Age Publishing.

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, *14*, 397–418.

Denver Evaluation Network. (2015). About. Retrieved from http://www.denverevaluationnetwork.org/about.html

Dierking, L. D., Ellenbogen, K. M., & Falk, J. H. (2004). In principle, in practice: Perspectives on a decade of museum learning research. *Science Education, 88*(S1), S1–S3. doi: 10.1002/sce.20021

Falk, J. H. (2000). Assessing the impact of museums. *Curator, 43*(1), 5–7.

Falk, J. H., & Dierking, L. (1992). *The museum experience*. Washington, DC: Whalesback Books.

Falk, J. H., & Dierking, L. D. (2000). *Learning from museums: Visitor experiences and the making of meaning*. Walnut Creek, CA: AltaMira Press.

Falk, J. H., & Sheppard, B. K. (2006). *Thriving in the knowledge age: New business models for museums and other cultural institutions*. Lanham, MD: AltaMira Press.

Fetterman, D. M., & Wandersman, A. (2005). *Empowerment evaluation principles in practice*. New York, NY: Guilford Press.

Gorman, A. K. (2007). Museum education assessment: Designing a framework. In P.

Villeneuve (Ed.), *From periphery to center: Art museum education in the 21st century* (pp. 206–212). Reston, VA: National Art Education Association.

Grack Nelson, A., & Callahan Schreiber, R. (2009). Participatory evaluation: A case study of involving stakeholders in the evaluation process. *Visitor Studies*, *12*(2), 199–213.

Gruber, D. D., & Hobbs, J. A. (2002). Historical analysis of assessment in art education. *Art Education*, *55*(6), 12–17.

Hooper-Greenhill, E. (2004). Measuring learning outcomes in museums, archives, and libraries: The learning impact research project (LIRP). *International Journal of Heritage Studies, 10*(2), 151–174.

Hein, G. E. (1998). *Learning in the museum*. London, UK: Routledge.

Hein, G. E. (2005). *The role of museums in society: Education and social action*. Retrieved from http://www.museoliitto.fi/seminaarit/museolehtoripaivat05/esitelmat/Finland-lec.1PC.pdf

Jacobsen, J. W. (2010). A Research vision for museums. *Curator*, *53*(3), 281–289.

Kansas City InfoZine. (2011). *Nelson-Atkins Museum of Art names education director* [Press release]. Retrieved from http://www.infozine.com/news/stories/op/storiesView/sid/49322

Kelly, L. (2004). Evaluation, research and communities of practice: Program evaluation in museums. *Archival Science*, *4*, 24–69 . doi 10.1007/s10502-005-6990-x

King, J. A, (2005). Participatory evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 292–297). London, UK: Sage.

King, J. A., & Volkov, B. (2005). A framework for building evaluation capacity based on the experiences of three organizations. *CURA Reporter*, *35*(3), 10–16.

Kirkpatrick, D. L. (2003). Evaluating training programs (2nd ed.). San Francisco, CA: Berrett-Koehler.

Korn, R. (2007). New directions in evaluation. In P. Villeneuve (Ed.), *From periphery to center: Art museum education in the 21st century* (pp. 213–218). Reston, VA: National Art Education Association.

Korn, R. (2008). Transforming museums —To what end? Proceedings: Transforming museums conference, Seattle, WA. Retrieved from: http://www.randikorn.com/docs/transforming_to_what_end.pdf

Kubarek, J. (2015). Building staff capacity to evaluate in museum education. *Journal of Museum Education*, *40*(1), 8–12.

Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, *33*(3), 307–338. doi: 10.1177/1098214011434608

Linn, R. L. (2000). Assessment and accountability. *Educational Researcher*, *29*(2), 4– 16.

Luke, J. J., & Ancelet, J. E. (2014). The role of evaluation in reimagining the art museum. *Journal of Museum Education*, *39*(2), 197–206.

Luke, J. J, & Knutson, K. (2010). Beyond science: Implications of the LSIE report for art museum education. *Curator*, *53*(2), 229–237.

MacDonald, B. (1974). Evaluation and the control of education. In B. MacDonald & R. Walker (Eds.), *Innovation, evaluation, research and the problem of control* (pp. 37–48). Norwich, UK: Centre for Applied Research in Education, University of East Anglia.

Owen, K., & Visscher, N. (2015). Museum-university collaborations to enhance evaluation capacity. *Journal of Museum Education*, *40*(1), 70–77.

Parsons, C. (2004). Let's evaluate programs. *Visitor Studies Today*, *7*(2), 2.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.

Patton, M. Q. (2015). *Qualitative evaluation and research methods* (4th ed.). Newbury Park, CA: Sage.

Preskill, H. (2011). Museum evaluation without borders: Four imperatives for making museum evaluation more relevant, credible, and useful. *Curator*, *54*(1), 93–100.

Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.

Rodríguez-Campos, L. (2005). *Collaborative evaluations: A step-by-step model for the evaluator*. Tamarac, FL: Llumina Press.

Rodríguez-Campos, L. (2012). Stakeholder involvement in evaluation: Three decades of the American Journal of Evaluation. *Journal of MultiDisciplinary Evaluation, 8*(17), 57–79.

Rowe, W. E., & Jacobs, N. F. (1998). Principles and practices of organizationally integrated evaluation. *Canadian Journal of Program Evaluation*, *13*, 115–138.

Schwandt, T. A. (2003). Back to the rough ground: Beyond theory to practice in evaluation. *Evaluation*, *9*(3), 353–364.

Schwandt, T. A. (2009). Globalizing influences in the western evaluation imaginary. In K. E. Ryan & J. B. Cousins (Eds.), *The Sage handbook of educational evaluation* (pp. 19-36). Thousand Oaks, CA: Sage.

Senge, P. (1990). *The fifth discipline: The art and practice of the learning organization*. New York, NY: Doubleday.

Shadish, W. R., Cook, T., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.

Sheppard, B. (2000). Do museums make a difference? Evaluating programs for social change. *Curator, 43*(1), 63–74.

Sheppard, B. (2010). Insistent questions in our learning age. *Journal of Museum Education, 35*(3), 217–228.

Steele-Inama, M. (2015). Building evaluation capacity as a network of museum professionals. *Journal of Museum Education*, *40*(1), 78–85.

Suchy, S. (2004). *Leading with passion: Change management in the twenty-first century museum*. Lanham, MD: AltaMira Press.

Thody, A. (2006). *Writing and presenting research*. London, UK: SAGE.

Worts, D. (2006). Measuring museum learning. A critical assessment framework. *Journal of Museum Education, 31*(1), 41–48.

York, P. (2005). *Learning as we go: Making evaluation work for everyone*. Briefing paper. Retrieved from http://www.tccgrp.com/pdfs/per_brief_lawg.pdf

---

[i] In this paper *program evaluation* refers to evaluating the intended and unintended consequences of participation in formal programs offered by museum education departments.
[ii] The umbrella term *museum* is used to refer to both art museums and public art galleries.

ⁱⁱⁱ *Visitor research* is the term used for the systematic collection and analysis of data to inform decisions about interpretive exhibits and programs within informal learning environments.