

# Notes on Distance Learning for Informal Settings

*White Paper #3: Identifying and measuring success*

July 31, 2023

**Prepared by:**  
Donnelley Hayde, M.A.  
Joe E. Heimlich, Ph.D.  
Gary Timko, Ph.D.

**Prepared for:**  
Indiana University



This project was completed with support from the National Science Foundation (#2115637).

**Recommended citation:** Hayde, D., Heimlich, J.E., Timko, G. (2023). *Notes on Distance Learning for Informal Settings, White Paper #3: Identifying and measuring success.* COSI's Center for Research and Evaluation.

## Overview

This literature review is funded in part by a grant from the National Science Foundation #1713567 and prepared for Indiana University as part of a larger project entitled *Data Visualization Literacy: Research and Tools that Advance Public Understanding of Scientific Data* (Dr. Katy Börner, Principal Investigator). This white paper is one component of a larger review. The larger review consists of three white papers, the first of which explores definitions, the history of distance education, and the technologies used for distance education over time. The second white paper explores the best practice considerations most widely employed in distance education. The present document focuses on the use of and strategies for applied evaluation of distance education programs. Together, the series of white papers aims to identify key elements of distance education across contexts, as well as the transferability of these approaches to informal science learning institutions.

## Introduction

As discussed in part 1 of this white paper series, a major concern related to distance education in the early days of the COVID-19 pandemic was the need to implement the infrastructure and practices of remote teaching and learning relatively quickly and universally amid a global catastrophe (Butcher, 2020). While both informal and formal learning settings quickly adapted (and sometimes exchanged) their actual programmatic offerings to provide some continuity (e.g., Higdon, 2020), systematic evaluation and reporting related to outcomes proved especially challenging, particularly as educators struggled to ensure that distance education was taking place at all (Ennes & Lee, 2021). Even as many educators entered this new context without significant training or experience in distance education, fewer still had existing capacity to measure its success. Meanwhile, amid concerns about learners “falling behind” socially and academically due to disruptions in in-person learning, the differences in the needs and goals of measurement in formal and informal distance education also seem to have been exacerbated. Where formal educators were forced to administer standardized tests that may or may not have reflected the contents and contexts of their immediate class experiences, informal educators faced significant pressure to provide positive leisure experiences that could generate revenue - and consequently had less cause and fewer opportunities to measure learning itself.

Despite these unique circumstances, the challenges and affordances of distance education emerged long before 2020, and they have continued to evolve even as in-person educational experiences have resumed. Acknowledging the specific pressures of COVID-19 and differences in learning contexts, the present document takes a long view of the existing scholarship on the evaluation of distance education in *both* formal and informal settings.

The authors acknowledge that peer-reviewed scholarship on this subject is unlikely to keep pace with real-world practice given the long time-step associated with academic publication and the immediacy of needs within the field (cf. Hardee & Duffin, 2015); by looking backward, we attempt to identify considerations for this work that are broadly applicable.

Rather than try to define best practices in evaluating distance education in informal learning settings, we have instead attempted to summarize historical context and perspective that can help frame these efforts in an intentional way that responds to the limited amount of scholarship on this very specific subject (Ennes & Lee, 2021). Furthermore, we recognize that beyond repositories for programmatic reporting (e.g., [informalscience.org](https://informalscience.org)), access to scholarly literature is relatively scarce for informal learning practitioners without academic library access. By synthesizing existing theoretical and practical work on the evaluation of distance education broadly, we hope to support informal learning practitioners in applying “lessons learned” to the changing landscape for this work, and particularly beyond the historically particular context of COVID-19.

## Method

This paper builds on the circular search strategy using Google Scholar initiated with the first paper in the series. Accordingly, we did not bound the literature within a time frame. As with the first white paper, the literature searches began with reading some of the handbooks on distance education, and articles and chapters reflecting on the history of the field. Notes were taken as to key terms used, who authored chapters (and in what areas), and key findings, conclusions, or implications offered with citations. These terms were then searched to find if there were additional concepts or ideas that should be added. The citations were explored to see what sources or authors were consistently used as references for different topics or referred to as authoritative. Also as before, this review represents a broad sweep and synthesis of the relevant literature, though its scope is limited to how evaluation is represented in scholarly writing about distance education. Rather than representing specific practices as ideal, the present white paper attempts to ground the practice of evaluation in particular considerations that can inform decision-making about evaluation of distance education programs.

## Why evaluate distance education?

In the context of distance education, evaluation can serve a variety of purposes, many of which are similar to those in in-person learning contexts. Possible purposes for measuring a program’s value and effectiveness might include justifying how resources are committed and used within the program, measuring the degree to which the program meets its goals, identifying ways to improve the program, and/or

determining whether the program should continue (Thompson & Irele, 2003). Meanwhile, evaluation can play an important role in identifying and documenting the unique considerations associated with different formats: as observed in parts 1 and 2 of this white paper series, changes to the learning space and the ways teachers and learners interact with it necessitate adaptations that leverage the specific affordances and acknowledge the specific limitations of the context.

Much of the literature on the evaluation of distance education focuses on how it can be used to inform program development and implementation. For example, Ruhe & Zumbo (2008) explain that evaluation is a useful tool for distance education because it supports continuous improvement in the context of change: in their focus area of formal learning, they propose it as a core element of a cycle, in which course evaluation follows one iteration of course development and implementation and precedes the next. In light of research that demonstrates distance education generally can be an effective approach to teaching and learning, Simonson (1997) also argues that the importance of evaluation to distance education lies in its utility for improving distance education and demonstrating the specific value of given distance education activities.

One argument for why evaluation may be *especially* important to distance education is that the platforms by which people teach and learn remotely are typified by rapid technological change (Ruhe & Zumbo, 2008). Thompson & Irele (2003) observe that throughout much of the history of distance education, evaluation was largely used as a vehicle to demonstrate that distance education was as valuable as traditional models of in-person education. In the context of online versions of distance education, they contend that the need for evaluation is more pronounced, in that online formats are particularly likely to contribute to high-impact changes within educational contexts; because of this, they also argue that it is particularly important that evaluation of distance education be appropriately rigorous (Thompson & Irele, 2003). In reviewing evaluation models specific to distance education (specifically, Flagg, 1990; Duning et al., 1993; Ehrmann, 1997b, 1999, 2001; Bates, 2000, and Cyr, 2001), Thompson & Irele (2003) particularly call attention to the dynamics of the learning space; while evaluation of distance education has much in common with evaluation of education generally, it is not sufficient to assume that the pedagogical and/or andragogical elements of distance education are the same as those of in-person learning experiences. Because of this, they contend that any comprehensive evaluation of distance education programs must do more than add assessment of basic technological function, accessibility, and user experiences to traditional content standards in order to be considered rigorous (Thompson & Irele, 2003). Relatedly, Lockee, Moore, & Burton (2002) point out that evaluation of distance education must not only examine individual components of a system, but also attempt to describe how the parts work together as a whole to support positive learning outcomes. To make sense of these issues, teams must begin by defining success within their specific distance education contexts.

## Identifying realistic, relevant goals

A crucial step in evaluating the success of a learning experience is articulating its goals, as well as the mechanisms by which those goals are to be achieved. In the context of distance education, learning goals often parallel those of in-person learning experiences, but they may be advanced through entirely different means and by learners with different opportunities and challenges. As Anastasiades (2003) observes, “Web-based learning designs must consider the nature of content, specific context, desired learning outcomes and characteristics of the learner” (p. 20). While these elements could be generalized as important to any learning environment, special attention to them is necessary when teaching strategies are implemented for different users and/or on different platforms.

One way to articulate goals is to consider what a high-quality distance learning experience looks like. Scholars of formal learning settings offer some guidance on defining quality in distance education, which often consists of bringing together best practices for teaching and best practices for usability and interface design (cf. Graham, et al., 2000; Parsons & Ryu, 2006). As mentioned in part 2 of this white paper series, the formats and contexts for teaching and learning remotely themselves have effects on learning; thus, quality depends on attention to these features of the experience, both in terms of their own functionality and in terms of how they are leveraged in a learning experience. In thinking about what this means for goals, Sherry (2003) describes quality as being affected by faculty- and institution-level priorities, but encourages practitioners to think about success as centered on learners, with specific attention to attitudes, competencies, applications, and impacts.

Within the literature on distance education in formal contexts, evaluation is itself also used to help define goals, through the process of needs assessment. For example, Watkins & Kaufman (2003) argue that needs assessment can be an important tool for evidence-based strategic planning related to distance education, with the premise that distance education should respond to clear needs and serve as a mechanism for meeting those needs (as opposed to being a goal in and of itself). Similarly, Thompson & Irele (2003) describe evaluation as something that, if used early in planning and development processes, can help provide evidence that investment in distance education is a justifiable investment of resources. Another way that evaluation can support meaningful targets for distance education is by contributing to the process of refining definitions of value and prioritizing specific forms of value. However, to do this, evaluation must be rigorous and credibly centered on questions of quality. To this point, the work of Hentea, Shea, & Pennington (2003) suggests that evaluation criteria for distance education must reflect important areas of student mastery to the same standards as those for in-person learning. Furthermore, the authors caution against superficial forms of assessment (e.g., measures of student satisfaction, rote memory tests, etc.), noting that these approaches can undermine the credibility of

distance education and actually serve to undermine student motivation (Hentea, Shea, & Pennington, 2003).

Outside of formal learning, definitions of success in distance education are less likely to reflect standardized curricula for learners at a particular age or skill level; instead, they reflect the particular learning needs of their contexts. For example, Roffe (2002) noted that in a corporate setting, areas of evaluative study might be those of particular interest in a for-profit setting: 1) the number of people going through a professional learning program, 2) efficiency of resource use, 3) the program's effectiveness (operationalized as including learner satisfaction, as well as changes in knowledge, behaviors, and productivity), and 4) the program's return on investment. Within institutions that specialize in informal learning (e.g. museums), evaluation of distance education similarly responds to particular needs and interests.

Importantly, remote engagement opportunities in the informal learning sector themselves reflect significant variation in format, ranging from digital programs and exhibitions hosted by institutions to entire museums that exist online, as well as a long tradition of physical resources meant for permanent distribution or loan (e.g., kits intended for individual or classroom use). Many distance education programs in informal learning began as a way to expand access to offerings at physical museums. For example, the Philadelphia Museum of Art's earliest distance learning programs leveraged digital telephone networks to "allow students to take field trips without leaving their seats" and to "give classroom teachers more flexibility to decide when and how to experience museums" (O'Leary, 2011). Within the general idea of remote learning, there are also informal learning resources that do not maintain a strict teacher-learner relationship: Styliani et al., (2009) describe "virtual museums" as digital collections made available both in-person via kiosks and remotely via websites, sometimes with a gamified or immersive component (e.g., virtual reality experiences), and generally without live human facilitation. These resources are sometimes leveraged to improve public access to collections objects and/or to preserve data about the objects.

When they interviewed informal science learning professionals with experience in distance learning about their perceptions of effective practices, Hardee & Duffin (2015) found that distance programming at informal learning organizations was generally conceived in terms of three key components: program design, program delivery, and a business model. Although the interviewees did not speak much to evaluation of their distance education programs, they frequently described programmatic success in terms of providing novel, appealing experiences (e.g., interactions with live animals, behind-the-scenes content, enrichment activities to be completed outside of synchronous programs, interactivity) and meeting educational standards. In addition, interviewees highlighted the importance of making efficient, budget-sensitive choices related to staffing and technology and considering business issues like marketing programs and balancing revenue against expenses.

In the context of little centralized planning or resources for distance education in informal learning (Hardee & Duffin, 2015), some museums have explored questions of scale. Through literature reviews for the National Air & Space Museum, Randi Korn & Associates, Inc. (2013) identified evaluation as an essential aspect of accountability for informal learning organizations interested in scaling up their programs to achieve greater impact. In addition, they recommend identifying the degree of alignment between the goals of a given program, those of its home organization, and those of potential partners and funders. Taken together, existing resources suggest that regardless of size, scope, or intended audiences, perhaps the most important element of defining success in informal distance education is identifying the purpose of the experience.

## Identifying what you need to know

After a distance education program's goals are articulated, the role of evaluation can vary depending on what kind of information is necessary to improving or reporting on the success of the program. Simonson (1997) highlights as possibilities Woodley & Kirkwood's (1986) six categories of measurement (activity, efficiency, outcomes, program aims, policy, and organizations) and the AEIOU approach, which attends to accountability, effectiveness, impact, organizational context, and unanticipated consequences (cf. Fortune & Keith, 1992; Sweeney, 1995; and Sorenson, 1996). Hew, et al. (2005) note the importance of naming the scale of evaluation questions and offer three possible levels of focus: the evaluation of whole programs, the evaluation of specific courses, and the evaluation of students' learning. In the case of students' learning, Phipps & Merisotis (1999) further specify student outcomes, student attitudes, and student satisfaction as three categories of measures typically used. Broadly speaking, the philosophical approach of the educational program also contributes to framing evaluation questions as related to learners. For example, Vrasidas (2000) compares objectivist and constructivist approaches to distance education, explaining that evaluation of a program with an objectivist approach may focus more on documenting learners' behavior change and cognitive understanding of a particular set of ideas. Meanwhile, the evaluation of a program with a more constructivist approach is more likely to involve documenting learners' growth and change in skills (e.g., problem-solving and knowledge construction). Accordingly, the choice of methods and the ways in which evaluation involves learners will vary depending on what kinds of data are considered most valuable and how such data will be used.

Another way of framing evaluation of distance education is via categories that concern themselves less with scale and more with key areas of implementation. Valcke & Leeuw (1998) identify five types of evaluation related to remote learning activities: 1) internal evaluation against specific performance indicators; 2) internal evaluation related to participant attitudes and perceptions, as well as staff performance; 3) external evaluation related to broader issues such as socio-cultural



environment, cost-benefit, feasibility, and networks; 4) the identification of interested/affected parties and their goals; and 5) issues related to the context of distance learning (e.g., selection and quality control of media and platforms). For the purpose of this synthesis, the remainder of this section is organized according to four common types of question in the evaluation of distance education.

## Question 1: Does it work for teachers?

For formative evaluation generally, functionality and usability are important areas of concern. In the context of distance education, functionality and usability frequently involve multiple dimensions, including various platforms for teaching and learning, practices for teaching and learning, and supplemental resources (activities, textbooks, etc.). To address any of these dimensions effectively, evaluation must clearly identify the roles of different operators within a programmatic system, along with how they relate to the goals of a given program. In other words, to define whether or how well something works for users, one must understand who counts as a user and what about their experience is important to measure.

### Teachers as learners

As distance education programs emerge, they often require some focused learning or adoption of practices on the part of teachers to ensure that the program itself is effective. Therefore, it can be important to the development of programs to evaluate teachers' own learning and perceptions. In the formal learning sphere, this kind of evaluation is sometimes implemented in the specific context of remote teacher education (i.e., distance education courses for pre-service teachers and/or continuing education for those already teaching), as opposed to within the evaluation of the teachers' own instruction (e.g., Krall, Straley, Shafer, & Osborn, 2009). Writing about remote teacher education, Perraton, Creed, & Robinson (2002) also describe the affordances of distance courses in terms of what one can easily evaluate. They contend that while it is fairly straightforward in distance education to gather data about what teachers know or understand, or even how they apply knowledge to their practice, it is much more challenging to directly authenticate teachers' actual performance remotely.

Meanwhile, distance (and sometimes hybrid) education offerings developed by informal learning organizations also sometimes include teacher training. The stated purpose of these training experiences is variable, and can range from improving teachers' knowledge and perceptions about specific content (e.g., Miele, Shanley, & Steiner, 2010; Steiner et al., 2016) to improving their skills related to specific teaching practices (e.g., Milligan, 2016; Smithsonian Organization and Audience Research, 2019). As in formal settings, the evaluation of these efforts typically focuses on what teachers learn and how they will apply it to their own practice.

## Teachers as users

When teachers are not positioned as learners in the context of evaluation, they may instead be considered in their critical role in program delivery. In this case, evaluation might focus on usability and implementation of resources for distance learning. For example, Clifton (2017) examined the outcomes of incorporating specific principles of learning design into a module for formal distance education; though the study leverages student data, a key element of the study is the comparison of student responses to those from developers and instructors. This comparison was critical to identifying a significant disconnect between the intentional design of the modules and what it was like to actually teach and learn with the module. Similarly, Perreault et al. (2002) undertook a study of business professors' experiences related to developing and teaching distance courses in order to better understand and address perceived barriers to implementation.

In informal learning settings, school-based teachers may be especially likely to be positioned as users due to their holding a customer-service provider relationship with the organization developing a distance education offering. Understanding and negotiating this relationship can be especially important to developing offerings that will appeal to and work for teachers. In these situations, evaluation may be limited to asking teachers their perceptions of a program's technical quality after the fact (e.g., Vitto, 2004), or it may reflect significant exploration of what outcomes are most important to teachers and what contextual features of a teacher's learning environment might affect those outcomes before the program begins (e.g., Tisdal, 2015). Other areas of study related to teachers as users can include ways of finding distance education opportunities and resources, how such opportunities and resources are integrated into classrooms, and what types of support for teachers might be helpful or necessary (Smithsonian Center for Learning and Digital Access with the School of Education at the University of California, Irvine, 2018).

In addition to school-based teachers, a less frequently discussed constituency within informal distance education is museum professionals. While museum professionals are usually those responsible for developing the distance education programs being described in evaluation reports, they also may represent a group for whom usability and/or contextual dynamics is especially important, as they often mediate designed learning experiences in similar ways to school-based teachers. As Sylaiou et al. (2008) suggest, involving these users early on in the design and testing of resources for distance education can identify critical aspects of technical development, user instructions and support, and implementation strategies.

## Question 2: Does it work for learners?

As developers of distance education offerings seek to build more empathy for learners into their designs (Matthews et al., 2017), formative studies focused on learners can provide important data about the "end-users" for distance education, or those for

whom the educational offering is planned. As Holmberg notes (1995), exploring students' experience, perceptions, and evidence of learning is important for understanding implementation and identifying how best to satisfy students' needs (190). As noted in the previous white papers, the context of learners' interactions with instructors, classmates, and classroom resources, including technology, can meaningfully affect the overall learning experience. Accordingly, evaluation involving how well distance education works for learners often explores the context of implementation, with special attention to the platforms and mechanisms through which learning happens.

Within formal learning settings, a particularly important consideration for design is what reasons learners have for participating in distance education. As noted in the first white paper of this series, distance education grew out of the reality that some adults have obligations that make traditional classroom experiences difficult to access. While the COVID-19 pandemic made the need for distance education more immediate for more people, persistent accessibility barriers to traditional classroom learning have made distance education salient well before and well after that historical moment. Since there are many potential delivery mechanisms for distance education, evaluation is sometimes used to gather feedback about how well a given implementation strategy addresses the need, whether by comparison to in-person courses or on its own terms (e.g., Puzzuoli, 1970; McConnell & Sharples, 1983; Coe & Elliott, 1999; Ligon, Markward, & Yegidis, 1999). Notably, some evaluation of distance education focuses on the distinct opportunities of distance education that differentiate it from traditional classroom learning experiences, particularly as new technologies emerge (Thompson & Irele, 2003). Areas of inquiry may include students' perceptions of their own learning, their interactions with instructors, the organization and pace of a course, instructional resources and methods, and/or technology used in a course (Puzzuoli, 1970; Coe & Elliott, 1999; Ligon et al., 1999; Gilroy et al., 2001; Schenker & Scadden, 2002; Abas, Lim, & Woo, 2009; Liu, 2012). Meanwhile, the accessibility and usability of technology used in distance education stands out as an especially common concern of evaluation in both formal settings (e.g., Schenker & Scadden, 2002; Motiwalla, 2007) and informal ones (e.g., Teather & Wilhelm, 1999; Sharples, 2000; Milligan et al., 2017; Graves & Wadman, 2017).

### Question 3: Does it convey what it intends?

Another major area of concern for evaluation of distance education is program effectiveness. Depending on the needs and intentions surrounding the program, studying a program's effectiveness may provide formative feedback or summative findings that document how much the program met its goals. In both formal and informal learning settings, content is a frequent focus for evaluation of this type, especially when desired learning outcomes for a program involve knowledge or understanding of specific concepts (see, e.g., Hershey, 1977; Kabat & Friedel, 1990; Buckley 2003; U.S. Department of Education, 2010; RK&A, 2019). Notably, measurement of outcomes related to content learning also appears in the literature in

explicit connection with non-formal distance education experiences, such as television programs that either stand on their own or complement a course of study (e.g., Anderson, 1998; Akhter, 2011).

In some cases, experiences may be more important to a distance education program's intended outcomes than specific learning content. In formal learning settings, this might mean evaluation focuses on something like a school's organizational climate (cf. Aluko & Shonubi, 2014) or elements of clinical practice (cf. Ligon et al., 1999). Experience-focused evaluation might also examine social aspects of the learning interaction (McCullough, 1997; Roblyer & Wiencke, 2003; Martens et al., 2007). Meanwhile, in informal learning settings like museums, evaluation might focus on learners' experience with novel settings or speakers, animals, collections objects, or specific approaches to teaching (e.g., Teather & Wilhelm, 1999; Ba & Keish, 2004; Morse et al., 2019).

#### Question 4: Whether and how should the program move forward?

Evaluation of distance education may also play an important role in informing organizational decision-making about the adoption, implementation, and/or refinement of programs. While any of the other questions described above might contribute to such decision-making, evaluation of distance education can also itself explore system-level criteria for effectiveness. (cf. Muilenburg & Berge, 2001; Lockee et al, 2002; Chao & Chen, 2009). This kind of evaluation is typically summative and particularly likely to be motivated by a need to justify the use of resources (including money, labor, and technological investment) on a given distance education program (Thompson & Irele, 2003).

### Identifying strategies and mechanisms for data collection

In naming some general principles for the evaluation of distance education, Simonson (1997) highlights the value of both quasi-experimental and naturalistic approaches to gathering and analyzing data and generally concludes that those seeking to do this type of work should use both quantitative and qualitative procedures. Ruhe & Zumbo (2008) practically detail end-to-end strategies for designing and implementing evaluation studies of distance education, and they similarly note that mixed-methods approaches provide comprehensive coverage of the dynamics involved in a distance education program. Sherry (2003) similarly notes that different areas and depths of focus for evaluation of distance education may be best addressed via different sources of data.

As discussed in the other parts of this white paper series, the various platforms used for distance education efforts often have unique affordances and challenges for teaching and learning. Accordingly, platforms also play a role in determining the

possibilities and limitations for the evaluation of distance education. As an example, Salomon (1984) compares the “amount of invested mental effort (AIME)” between watching a television program and reading a printed text (p. 647) and observes that learners’ perceptions of different mediums seem to play a role in what they take away from a learning experience. Similarly, Passerini & Granger (2000) contend that the specific types of peer and instructor interactions enabled by the internet have been critical to moving toward more constructivist approaches to distance education (p. 4). Each of these cases describes interactions with platforms that fundamentally change how the learning might be described and how learning can be attributed to an intervention. Just as educators must consider the appropriateness of their teaching strategies for a specific context, evaluators must consider what ways of gathering data are most useful for that context and best reflect the sources of evidence necessary to answer their questions. Moreover, distance education programs may present unique ethical considerations for authenticating and protecting student identities and identifying how data are collected (Hentea et al., 2003; Traxler, 2007).

The practicalities of evaluating distance education also change depending on learners’ physical locations and the types of tools they can access. As Baker (2003) explains, evaluation must proactively adapt to the context of the program, lest factors like physical distance, indirect communication, and lack of controls weaken study credibility. While some platforms for distance education, such as computers or direct paper mail, might mean it is easiest to conduct evaluation in the same format as the learning experience (e.g., Galloway, 2005), others may provide opportunities to leverage multiple types of data. In the case of telephone-based distance education programs, Olgren (1997) recommends using print materials mailed directly to learners and audio recordings of program activities. A particularly thorough example of managing data collection across platforms and at various time points can be found in Anastasiades (2003), which combined questionnaires and live discussion at various stages and among different participants in a hybrid distance education program. Although both examples reflect technological solutions that may now be rarer in the age of ubiquitous digital resources, both provide a useful demonstration of how to work with available technologies, as their combination of approaches not only provides multiple sources of evidence but also leverages both synchronous and asynchronous interactions.

Within the specific domain of digital, internet-based platforms, there may be specific opportunities to leverage technology in service of evaluation. For example, it is sometimes possible to embed analytics into program resources (e.g., Interactive Educational Systems Design, Inc., 2017), to ensure that key interactions are documented via chat communications (Munzer, 2003), or to build the demonstration of skills into a game (Bidarra et al., 2011). Documentation of how much time learners spend on computers and how they tend and prefer to use technology may also itself be valuable contextual data to understanding how particular platforms interact with learning (Lei & Zhao, 2007; Smithsonian Center for Learning and Digital Access and Navigation North Learning Solutions, 2017). In general, Holmberg (1995) provides a useful touchstone for those tasked with evaluating distance education, noting that the

unique context “opens up a number of possibilities...and it does so in ways that are different from those of conventional education” (p. 206). Thus, effective evaluation must consider and leverage context accordingly.

## Implications for informal learning practitioners

In general, evaluation of teaching and learning in distance education can vary significantly depending the perceived goal of a program (e.g., knowledge acquisition vs. social or leisure experiences), what elements of the program represent institutional questions or persistent challenges (e.g., usability, learner awareness of key messages, etc.), and the practical affordances and limitations of platforms. For those tasked with measuring the success of distance education, this means the first and most critical step of study design is understanding the operating context and the purpose for the study. As the design of an evaluation proceeds, key priorities then include identifying what success means in context, then asking questions and using methods that reflect the core goals, orientations, and feasibly accessible data sources associated with a given program. While formal and informal learning settings may specify different resources and goals, and perhaps even different ways of thinking about audiences and learning, these basic process elements appear to be consistent across contexts.

By leveraging these high-level commonalities in process, informal learning professionals can conduct high quality evaluation of distance education that reflects a substantial amount of consideration within scholarly literature. Meanwhile, the frequent adaptation in formal distance education of methods common to evaluation of in-person courses (e.g., questionnaires, interviews, skills testing, etc.), suggests that evaluation of distance education in informal learning settings might similarly leverage its common “tools of the trade” used to measure the success of in-person programming, with mechanical adaptations to specific platforms where appropriate. Ultimately, successful, credible evaluation of distance education, regardless of context, means gathering internally meaningful evidence that reflect operational goals; for informal learning in particular, this may mean advocating for intentional, well-articulated planning of distance education efforts, as well as identifying strategies for measurement that honor experiences as important goals in their own right.

## References

- Abas, Z. W., Lim, T. S. K., & Woo, T. K. (2009). Mobile learning initiative through SMS: A formative evaluation. *ASEAN Journal of Open and Distance Learning*, 1(1), 49-58.
- Aluko, F. R., & Shonubi, O. K. (2014). Going beyond Kirkpatrick's Training Evaluation Model: The role of workplace factors in distance learning transfer. *Africa Education Review*, 11(4), 638-657.
- Anastasiades, P. S. (2003). Distance learning in elementary schools in Cyprus: the evaluation methodology and results. *Computers & Education*, 40(1), 17-40.
- Ba, H. & Keisch, D. (2004). *Bridging the gap between formal and informal learning: Evaluating the SeaTrek distance learning project*. Center for Children & Technology. <http://www.cct.edc.org/publications/bridging-gap-between-formal-and-informal-learning-evaluating-seatrek-distance-learning>
- Baker, R. K. (2003). A framework for design and evaluation of internet-based distance learning courses: Phase one-Framework justification design and evaluation. *Online Journal of Distance Learning Administration*, 6(2), 43-51.
- Bidarra, J., Rothschild, M., & Squire, K. (2011). Games and simulations in distance learning: The AIDLET model. In M. M. Cruz-Cunha, V. H. Varvalho, & P. Tavares (Eds.), *Computer Games as educational and management tools: Uses and approaches* (pp. 67-85). IGI Global.
- Buckley K. M. (2003). Evaluation of classroom-based, Web-enhanced, and Web-based distance learning nutrition courses for undergraduate nursing. *The Journal of Nursing Education*, 42(8), 367-370.
- Butcher, J. (2020, March 25). *Public-private virtual-school partnerships and federal flexibility for schools during COVID-19*. Special Edition Policy Brief. <http://dx.doi.org/10.2139/ssrn.3564504>
- Chao, R. J., & Chen, Y. H. (2009). Evaluation of the criteria and effectiveness of distance e-learning with consistent fuzzy preference relations. *Expert Systems with Applications*, 36(7), 10657-10662.
- Clifton, G. (2017). An evaluation of the impact of “learning design” on the distance learning and teaching experience. *International Review of Research in Open and Distributed Learning: IRRODL*, 18(5), 277-286.

- Coe, J. A. R., & Elliott, D. (1999). An evaluation of teaching direct practice courses in a distance education program for rural settings. *Journal of Social Work Education, 35*(3), 353-365.
- Ennes, M. & Lee, I.N. (2021). Distance learning in museums: A review of the literature. *International Review of Research in Open and Distributed Learning, 22*(3), 162-181.
- Galloway, D. L. (2005). Evaluation distance delivery and e-learning: Is Kirkpatrick's model relevant? *Performance Improvement, 44*(4), 21-27.
- Gilroy, P., Long, P., Rangecroft, M., & Tricker, T. (2001). Evaluation and the invisible student: theories, practice and problems in evaluating distance education provision. *Quality Assurance in Education, 9*(1), 14-22.
- Graham, C., Cagiltay, K., Craner, J., Lim, B. R., & Duffy, T. M. (2000). *Teaching in a web based distance learning environment* (Center for Research on Learning and Technology Technical Report No. 13-00).
- Graves, A., & Wadman, M. (2017). *Early learning with the Smithsonian Learning Lab*. Learning Lab Research. [https://learninglab.si.edu/cabinet/file/5809dc91-2f08-4025-b68b-5c1361322906/Smithsonian-SCLDA\\_EarlyLearning.pdf](https://learninglab.si.edu/cabinet/file/5809dc91-2f08-4025-b68b-5c1361322906/Smithsonian-SCLDA_EarlyLearning.pdf)
- Hardee, C., & Duffin, M. (2015). *Summary of evaluation findings: Digital programming in informal science settings, current trends and practices*. Retrieved from PEERassociates.net.
- Hentea, M., Shea, M. J., & Pennington, L. (2003, October). A perspective on fulfilling the expectations of distance education. In *Proceedings of the 4th conference on Information technology curriculum* (pp. 160-167).
- Hershey, M. (1977). Telephone Instruction: an Alternative Educational Delivery System for Teacher in-Service. *Gifted Child Quarterly, 21*(2), 213-217.
- Hew, K. F., Liu, S., Martinez, R., Bonk, C. J., & Lee, J. Y. (2005, October). Online Education Evaluation: What Should We Evaluate?. In *Association for Educational Communications and Technology Annual Meeting*.
- Higdon, D., Goldstone, J, Meads, C., Gabriel, K., Jepson, T., & Mathison, C. (2020). *School in the Park final report 2019-20*. [Technical report for Price Philanthropies.] [https://drive.google.com/file/d/14Sg2g7Ha7065v1pxpW2c9cy\\_JTbAXMMW/view](https://drive.google.com/file/d/14Sg2g7Ha7065v1pxpW2c9cy_JTbAXMMW/view)
- Holmberg, B. (1995). *Theory and practice of distance education* (2nd ed.). Routledge.



- Interactive Educational Systems Design, Inc. (2017). *Integrating authentic digital resources in support of deep, meaningful learning*.  
[https://learninglab.si.edu/cabinet/file/66dd07f4-9cc3-43e8-9e33-e688d9292c97/SCLDA-WP-DeepLearning\\_01172017.pdf](https://learninglab.si.edu/cabinet/file/66dd07f4-9cc3-43e8-9e33-e688d9292c97/SCLDA-WP-DeepLearning_01172017.pdf)
- Kabat, E. J. & Friedel, J. (1990). *The development, pilot-testing, and dissemination of a comprehensive evaluation model for assessing the effectiveness of a two-way interactive distance learning system*. [Technical report for Eastern Iowa Community College District.]
- Krall, R. M., Straley, J. P., Shafer, S. A., & Osborn, J. L. (2009). Hands-on at a distance: Evaluation of a temperature and heat distance learning course. *Journal of Science Education and Technology*, 18, 173-186. DOI 10.1007/s10956-008-9142-2
- Lei, J., & Zhao, Y. (2007). Technology uses and student achievement: A longitudinal study. *Computers and Education*, 49, 284-296. doi:10.1016/j.compedu.2005.06.013
- Ligon, J., Markward, M. J., & Yegidis, B. L. (1999). Comparing student evaluations of distance learning and standard classroom courses in graduate social work education. *Journal of Teaching in Social Work*, 19(1/2), 21-29. DOI: 10.1300/J067v19n01\_03
- Liu, O. L. (2012) Student evaluation of instruction: In the new paradigm of distance education. *Research in Higher Education*, 53, 471-486. DOI 10.1007/s11162-011-9236-1
- Lockee, B., Moore, M., & Burton, J. (2002). Measuring success: evaluation strategies for distance education. *Educase Quarterly*, 1, 20-26.
- Martens, R., Bastiaens, T., & Kirschner, P. A. (2007). New learning design in distance education: The impact on student perception and motivation. *Distance Education*, 28(1), 81-93.
- McCullough, K. H. (1997). Participatory evaluation in distance learning. *Open Learning: The Journal of Open, Distance and e-Learning*, 12(1), 24-30. DOI: 10.1080/0268051970120104
- Matthews, M. T., Williams, G. S., Yanchar, S. C., & McDonald, J. K. (2017). Empathy in distance learning design practice. *Tech Trends*, 61, 486-493.
- Miele, E., Shanley, D., Steiner, R. V. (2010). Online teacher education: A formal informal partnership between Brooklyn College and the American Museum of Natural History. *The New Educator*, 6, 247-264.

- Milligan, D. (2016, January 15). *Moving from downloads to uploads: Toward an understanding of the curricular implications of access to large scale digitized museum collections on the professional practice of K-12 classroom educators*. MW2016: Museums and the Web 2016. <https://mw2016.museumsandtheweb.com/paper/moving-from-downloads-to-uploads-understanding-curricular-implications-of-access-to-large-scale-digitized-museum-collections-on-the-professional-practice-of-k-12-classroom-educators/>
- Milligan, D., Wadman, N, & Ausman, B. (2017, February 1). *Discovering, creating, and sharing digital museum resources: a methodology for understanding the needs and behaviors of student users*. MW17: MW 2017. <https://mw17.mwconf.org/paper/discovering-creating-and-sharing-digital-museum-resources-a-methodology-for-understanding-the-needs-and-behaviors-of-student-users/>
- Morse, C., Koenig, V., Lallemand, C., & Wieneke, L. (2019, January 16). *Art in Rich-Prospect: Evaluating Next-Generation User Interfaces for Cultural Heritage*. MW19: MW 2019. <https://mw19.mwconf.org/paper/art-in-rich-prospect-evaluating-next-generation-user-interfaces-for-cultural-heritage-2/>
- Motiwalla, L. F. (2007). Mobile learning: A framework and evaluation. *Computers and Education*, 49, 581-596.
- Muilenburg, L. & Berge, Z. L. (2001). Barriers to distance education: A factor-analytic study. *American Journal of Distance Education*, 15(2), 7-22. DOI: 10.1080/08923640109527081
- Munzer, S. (2003). An evaluation of synchronous co-operative distance learning in the field: The importance of instructional design. *Educational Media International*, 40(1/2), 91-100. DOI: 10.1080/0952398032000092143
- O'Leary, L. (2011). Insights on a museum's distance learning program. *Journal of Museum Education*, 36(3), 241-248.
- Olgren, C. H. (1997). Teaching by telephone. *New Directions for Teaching and Learning*, 71, 59-66.
- Parsons, D., & Ryu, H. (2006). A framework for assessing the quality of mobile learning. Conference: 11th International Conference for Process Improvement, Research and Education (INSPIRE).
- Passerini, K., & Granger, M. J. (2000). A developmental model for distance learning using the internet. *Computers and Education*, 34, 1-15.

- Perraton, H., Creed, C., & Robinson, B. (2002). *Teacher education guidelines: Using open and distance learning. Technology, Curriculum, Cost, Evaluation*. UNESCO, Higher Education Division, Teacher Education Section.
- Perreault, H., Waldman, L., Alexander, M., and Zhao, J. (2002). Overcoming barriers to successful delivery of distance-learning courses. *Journal of Education for Business*, 77(6), 313-318. DOI: 10.1080/08832320209599681
- Phipps, R., & Merisotis, J. (1999). *What's the difference?: A review of contemporary research on effectiveness of distance learning in higher education*. The Institute for Higher Education Policy.
- Puzzuoli, D. (1970). *A study of teaching university extension classes by tele-lecture*. West Virginia University.
- Randi Korn & Associates, Inc. (2013). *Annotated Bibliography: Scaling Up and Distance Education*. [Technical report for the National Air & Space Museum].
- RK&A (2019). *Summative evaluation: Why Engineering? Program*. [Technical report for the Building Museum.] <https://www.informalscience.org/summative-evaluation-why-engineering-program>
- Roblyer, M. D., & Wiencke, W. R. (2003). Design and use of a rubric to assess and encourage interactive qualities in distance courses. *The American Journal of Distance Education*, 17(2), 77-98. DOI: 10.1207/S15389286AJDE1702\_2
- Roffe, I. (2002). E-learning: Engagement, enhancement and execution. *Quality Assurance in Education*, 10(1), 40-50.
- Ruhe, V., & Zumbo, B. D. (2008). *Evaluation in distance education and e-learning: The unfolding model*. Guilford Press.
- Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attitudes. *Journal of Educational Psychology*, 78(4), 647-658.
- Schenker, K. T., & Scadden, L. A. (2002). The design of accessible distance education environments that use collaborative learning. *Information Technology and Disabilities Journal*, 8(1),
- Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. *Computers & Education*, 34, 177-1
- Sherry, A. C. (2003). Quality and its measurement in distance education. In M. G. Moore & W. G. Anderson (Eds.), *Handbook of distance education* (pp. 435-459). Lawrence Erlbaum Associates, Inc.

- Simonson, M. R. (1997). Evaluating teaching and learning at a distance. *New Directions for Teaching and Learning*, 71, 87-94.
- Smithsonian Center for Learning and Digital Access and Navigation North Learning Solutions. (2017). *Technical Specifications for Future Development of the Smithsonian Learning Lab*. [Technical report for the Smithsonian Center for Learning and Digital Access.]
- Smithsonian Center for Learning and Digital Access with the School of Education at the University of California, Irvine (2018). *Curation of Digital Museum Content: Teachers Discover, Create, and Share in the Smithsonian Learning Lab*. [Technical report for the Smithsonian Center for Learning and Digital Access.] <http://s.si.edu/CurationofDigitalMuseumContent>
- Smithsonian Organization and Audience Research (2019). *An evaluation of "Curating Digital Museum Content with Visible Thinking Routines:" A training conducted by the Smithsonian Center for Learning and Digital Access*. [Technical report for the Smithsonian Center for Learning and Digital Access.] [https://learninglab.si.edu/cabinet/file/87241dbe-7a83-4f1f-9427-c80eed672152/Smithsonian-SCLDA\\_CuratingDigitalMuseumContentwithVTR.pdf](https://learninglab.si.edu/cabinet/file/87241dbe-7a83-4f1f-9427-c80eed672152/Smithsonian-SCLDA_CuratingDigitalMuseumContentwithVTR.pdf)
- Steiner, R., Doherty, P., Payo, R., Randle, D., & Stokes, L. (2016, February 4). *A multi-institutional model of blended teacher professional development*. MW2016: Museums and the Web 2016. <https://mw2016.museumsandtheweb.com/paper/the-blended-brain-inter-institutional-teacher-professional-development-among-three-informal-science-institutions/>
- Styliani, S., Fotis, L., Kostas, K., & Petros, P. (2009). Virtual museums, a survey and some issues for consideration. *Journal of cultural heritage*, 10(4), 520-528.
- Teather, L., & Wilhelm, K. (1999). Web musing: evaluating museums on the web from learning theory to methodology. *Museums and the Web 1999*. <https://www.archimuse.com/mw99/papers/teather/teather.html>
- Thompson, M. M., & Irele, M. E. (2003). Evaluating distance education programs. In M. G. Moore & W. G. Anderson (Eds.), *Handbook of distance education* (pp. 567-584). Lawrence Erlbaum Associates, Inc.
- Tisdal, C. E. (2015). Distance learning programs at the Saint Louis Zoo: Focus groups with teachers of K-12 students with disabilities. [Technical report for the Saint Louis Zoo.] <https://www.informalscience.org/distance-learning-programs-saint-louis-zoo>

- Traxler, J. (2007). Defining, discussing, and evaluating mobile learning: The moving finger writes and having writ.... *International Review of Research in Open and Distance Learning*, 8(2), 1-12.
- U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2010). Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. <https://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>
- Valcke, M. M., & Leeuw, F. L. (1998) *Evaluating digital distance learning programs and activities: Studies, practices, and recommendations*. World Bank Institute of the World Bank.
- Vitto, C.S. (2004). *If we build it, will they come? A year of testing distance learning using the web*. Museums and the Web 2004. <https://www.archimuse.com/mw2004/papers/vitto/vitto.html>
- Vrasidas, C. (2000). Constructivism versus objectivism: Implications for interaction, course design, and evaluation in distance education. *International Journal of Educational Telecommunications*, 6(4), 339-362.
- Watkins, R., & Kaufman, R. (2003). Strategic planning for distance education. In M. G. Moore & W. G. Anderson (Eds.), *Handbook of distance education* (pp. 507-517). Lawrence Erlbaum Associates, Inc.