Documenting and Assessing Learning in Informal and Media-Rich Environments:

A Report to the MacArthur Foundation

December, 2012

Documenting and Assessing Learning in Informal and Media-Rich Environments:

A Report to the MacArthur Foundation

December, 2012

Prepared by:

Jay Lemke, UC San Diego
Robert Lecusay, UC San Diego
Mike Cole, UC San Diego
Vera Michalchik, SRI International

Contents

Executive Summary 3
Introduction 6
Review of the Literature 13
Summaries of the Expert Meetings 53
Conclusions and Recommendations 62
References 70
Appendix A: Expert Meeting Participants 75
Appendix B: Bibliography 77
Appendix C: Online Resources: Assessment, Funding, Research 102

Exhibits

Exhibit ES-1. Outcomes-by-Levels Model for Documentation and Assessment 4
Exhibit 1. Outcomes-by-Levels Model for Documentation and Assessment 8
Exhibit 2. Clusters of Informal Learning Outcomes by Level, with Examples 10
Exhibit 3. Screen shot of the Scratch Programming interface. 20
Exhibit 4. Excerpt of graphic representation of one of Randy's digital stories 23
Exhibit 5. Features of online social learning networks (SNL) and their corresponding forms of data....25
Exhibit 6. Definitions of actions and concepts Nacu et al (2012) determined relevant for addressing their research questions 26
Exhibit 7. From Nacu et. al. (2012) 27
Exhibit 8. Simplified Competency Model for the commercially available role playing game *Oblivion*. ...39
Exhibit 9. From left to right: the pull-down destination menu, and two of the 30 destinations where players can visit and interact, the beach and the moon. From Fields and Kafai (2007). 45
Exhibit 10. Network graphs from different slices of the play history for one player in *Digital Zoo*.....49
Exhibit 11. Quest to Learn Objectives Model.....51
Exhibit 12. Three competency models for student assessment at Quest to Learn.....52

Executive Summary

The aim of this project has been to provide a review of relevant research on documenting and assessing learning in informal educational activities and to provide a set of related recommendations.

In 2011-12, the project team convened a series of 3 meetings of cumulatively 25 experts to provide advice and input to the research review and critical analysis of relevant concepts, categories, and approaches. The results of these wide-ranging discussions are summarized in this report and were highly influential in formulating our recommendations.

The research review generated an extensive bibliography from which we selected for description and analysis a subset of studies and projects to illustrate both the diversity of approaches to assessment of learning in informal activities and good assessment practices. The review and this report focus on:

- Learning in after-school programs and community centers
- Short-term, focused out-of-school activities
- Activities in informal learning institutions (museums, aquariums, zoos, etc.)
- Computer-based and online activities

We find that “informal learning” is a very broad category, but define it for our purposes as consisting of organized activities in face-to-face or online settings other than formal instruction in which a number of the following features are especially salient:

- Voluntary participation
- Relatively equitable power relations in negotiating goals and means
- Enjoyment of the learning activity for its own sake
- Intense engagement with tasks
- Flexibility of goals and in re-purposing resources
- Unpredictability of some significant learning outcomes
- Improvisation and innovation within and concerning the activity

Regarding documentation and assessment of such activities, our principal finding and recommendation is that ***the scope of valued learning outcomes be broadened to include social, emotional, and developmental outcomes as well as school-sanctioned knowledge, and to include learning by groups and whole projects as well as by individuals.*** In particular, we note that many of the valued learning outcomes reported are not ones that were predictable or aimed for at the start of projects.

We propose a synthesis of assessment approaches that encompasses at least ten general types of valued outcomes, to be assessed at each of three organizational levels (project, group, individual). We find that

outcomes at each level are frequently influenced by those at other levels, as when groups learn how to better support the learning of individual members and when projects learn how to better enable participants to identify and achieve their own learning goals.

Exhibit ES-1. Outcomes-by-Levels Model for Documentation and Assessment

Level of Analysis

Outcomes	Project/Community	Group	Individual
Social-Emotional-Identity Development	Developing social-emotional climate; community/project ethos, goals, local culture; system of roles and niches	Mutual support, challenge, inspiration; joint enjoyment and engagement	Comfort and sense of agency in domain; engagement; long-term interest and persistence vs. obstacles and frustration
Cognitive-Academic (Know-how	Developing strategies for organizing and distributing know-how; work practices, division of labor	Shared, distributed know-how; collective intelligence; dialogue and cooperation skills; explanation skills	Knowing how to go forward in the domain; Knowing how to mobilize and integrate know-how across domains

Among the specific outcomes that should be included in assessment, where relevant are:

- Increased know-how and know-who
- Persistence and resilience in the face of obstacles, setbacks, and frustration
- Improved ability to successfully guide one's own learning
- Improved ability to complete tasks cooperatively with others
- Production of improved and high-quality products
- Increased comfort with and sense of agency in a domain
- Improved ability to assess the credibility of claims and information sources
- Assisting others to achieve valued learning outcomes
- Widened range of mobilization of learning across tasks, domains, and settings
- Increased emotional maturity and productive use of affective sensibilities
- Continued development and application of learning over more than one year

Overall, we believe the unit of analysis for assessment should be a system of activities and practices over time and on multiple time-scales from hours to months or years.

Assessments which are made only at the end of a project are not likely to be sensitive to learning about how to do things better, nor can they identify the processes by which improvements came to be made. Assessments based only on the initial goals of a project may not capture other valued learning outcomes, which emerged unpredictably over time. Documentation and assessment that only sample outcomes on short time-scales (e.g. by

standardized testing) cannot assess practices that inherently take much longer to be enacted (e.g. designing a product, refining an experiment).

Effective assessment we find:

- Takes into account the social and cultural context of the communities and institutions supporting and constraining the activities being assessed, including community buy-in and prospects for long-term sustainability.
- Adopts a longitudinal design, following changes and gains from the inception of a project and documenting the processes by which desirable outcomes come about and not the outcomes alone.
- Adopts a broad definition of knowledge and learning which includes know-how and know-who as well as know-that, and which recognizes that groups and projects as well as individuals learn to do better.
- Aims to generate information that will be of practical use to individual learners, groups, projects, and organizations not only in identifying outcomes but also in planning for future improvements.

We identify a number of **promising directions** for the future of documentation and assessment of informal learning activities, including:

- Projects that expand the numbers and kinds of mentors and re-think their roles
- Computer-assisted learning games that unobtrusively document progress
- Data-mining and machine-learning analysis of extensive data archives
- Crowd-sourcing of assessment (e.g. open badge systems)
- Agent-based modeling techniques for simulating systems of practices with different outcomes
- Ethnographic documentation of naturally occurring activity offline and online and documentation of activity in richly-instrumented spaces

We find that the validity of **standardized testing** measures has not been persuasively established for informal learning activities, in large part because these activities do not have fixed, predictable curricular outcomes and because the valued outcomes they do have are unpredictable, not ones for which valid standard measures exist, or for which such measures cannot be used in the informal learning context without being disruptive of other primary goals and commitments (e.g. enjoyment, fostering creative production, maintaining a free-choice community, etc.).

We conclude with a number of more specific recommendations consistent with these general findings (p. 60).

Introduction

In 2010, the authors of this report were asked to review the relevant literature and convene a series of expert meetings to make recommendations regarding the state of the art and outstanding challenges in documenting and assessing learning in informal and media-rich environments.

For some years now, efforts such as the MacArthur Foundation's Digital Media and Learning initiative have supported the development of a range of educational activities, media, and environments outside the classroom and its formal curriculum. Other large-scale efforts, such as the National Science Foundation–supported LIFE Center (Learning in Formal and Informal Environments), have emphasized the complementarity of formal and informal learning experiences and the potential for educational reform to benefit from knowledge gained in the study of learning outside school. In a similar vein, the National Research Council produced a consensus report reviewing the knowledge base regarding learning in non-school environments (Bell, Lewenstein, Shouse, & Feder, 2009). Across these efforts, there is agreement that the success and scaling up of such initiatives depends on our ability to effectively document and assess what works and what doesn't, where, when, why, and how, in informal learning.

This report summarizes an extensive review of the literature on assessment of learning in such informal settings as after-school programs, museums and science centers, community-based organizations, and online communities. In addition, we convened three expert meetings involving a total of 25 participants to discuss key issues, identify successful approaches and outstanding challenges, and review summaries of prior meetings in the series. Our aim is twofold: first, to offer to those who design and assess informal learning programs *a model of good assessment practice,* a toolkit of methods and approaches, and pointers to the relevant literature; and second, to offer to the Foundation and other supporters and funders of projects *recommendations* concerning good practice in project assessment and identifiable needs for developing improved assessment techniques.

The members of our expert panels strongly urged us to deal with fundamental questions such as the purposes of assessment and the kinds of valued outcomes that should be considered. From discussions with the panel members and analysis of the research literature, as well as our own experience and judgment, we constructed a basic assessment model that encompasses at least ten general types of valued outcomes, to be assessed in terms of learning at the project, group, and individual levels. Not all levels or all outcome types will be equally relevant to every project, but we strongly believe that all assessment designs need to begin by considering a conceptual model at least as comprehensive as what we propose here.

This is particularly important because the valued outcomes of informal learning tend to be less predictable and much more diverse than those of formal education. Formal education is designed to strongly direct learning into particular channels and produce outcomes that are specifiable in advance and uniform across students. Informal learning experiences build on the diverse interests and curiosity of learners and support their self-motivated

inquiries. The valued outcomes of informal learning are often particularly rich in contributions to social and emotional development, to identity and motivation, to developing skills of collaboration and mutual support, and to persistence in the face of obstacles and inquiry on timescales of weeks, months, and even years. Informal learning activities also often result in products and accomplishments of which students are justly proud and for which product-appropriate measures of quality are needed.

In the remainder of this introduction, we will display our Outcomes-by-Levels Model for comprehensive assessment and briefly provide some definitions, distinctions, and principles as a general framework for what follows. We will then provide a review of selected and representative research studies and project reports in order to illustrate a wide range of useful techniques for documenting and assessing informal learning across varied settings and to identify issues and challenges in the field. Finally, we will provide our overall conclusions and recommendations.

Outcomes and Levels

It was universally agreed in our expert panels and extensively illustrated in the research literature that simple declarative knowledge is only one valued outcome of learning and is too often over-emphasized in assessment designs to the exclusion or marginalization of other equally important outcomes. Likewise, assessment designs too often focus only on outcomes for individual learners and neglect group-level learning and project-level or organization-level learning. Documentation and assessment need to be able to show how whole projects and supporting organizations learned to do better, or didn't. The kinds of documentation and data of value for organizational-level improvement are not limited to those that document individual learning.

Even individual learning is not simply a matter of declarative domain-specific knowledge. As an aspect of human development—at the individual, group, or organizational levels—the learning that matters is learning that is used. This type of learning plays a role in constructive activities: from posing questions to solving problems, from organizing a group to building a simulation model, from exploring a riverbank to producing a video documentary. In all these cases what matters is *know-how*, and *know-that* matters only insofar as it is mobilized in practice.

Activities of practical value usually require interaction and collaboration with other people. *Know-who* is as important as know-how to getting things done. Social networking and coming to understand who is good at what and how a group of particular people can work together effectively is an essential outcome of learning.

Nothing of value gets undertaken unless people are motivated to act and feel comfortable with the domains of know-how and the kinds of collaboration with others needed to get things done. A key outcome of learning is the development of identification with ideals, goals, groups, tools, media, genres, and styles that constitute our changing identities and motivations for action. Equally important is our social-emotional development in learning how to use our feelings, our emotional relations to others, and our emotional reactions to events for constructive purposes.

Collaborative groups learn, develop, and change over time. Membership may change; agreed goals, processes of interaction, interpersonal feelings, agreed procedures, and informal ways of doing things all change. In many cases they change adaptively so that the goals of the group are more effectively pursued. Just as individuals learn how to better function in collaborative groups, so groups learn how to make better use of the contributions of individual members. Or they don't.

Whole projects, online communities, and larger organizations also learn, change, and adapt. Or they don't. Documenting and assessing organizational learning is equally as important as assessing group and individual learning and development. It is likely, though not well understood, that learning processes at these three levels are linked and that we cannot expect to understand why learning was successful or unsuccessful at any one of these levels unless we also have data about learning at the other two.

From these and similar considerations we developed the following basic *Outcomes-by-Levels Model for Documentation and Assessment:*

Exhibit 1. Outcomes-by-Levels Model for Documentation and Assessment

Level of Analysis

Outcomes	Project/Community	Group	Individual
Social-Emotional-Identity Development	Developing Social-Emotional Climate; Community/Project Ethos, Goals, Local Culture; System of Roles & Niches	Mutual support, challenge, inspiration; Joint enjoyment & engagement	Comfort & sense of agency in domain; Engagement; Long-term interest and persistence vs. obstacles & frustration
Cognitive-Academic (Know-how)	Developing strategies for organizing and distributing know-how; Work practices, division of labor	Shared, distributed know-how; Collective intelligence; Dialogue and cooperation skills, explanation skills	Knowing how to go forward in the domain; Knowing how to mobilize and integrate know-how across domains

In addition to this basic outcome-types by levels matrix, we also need to emphasize the importance of taking into account in assessment design the incorporation of relevant knowledge about the history of the project, the community, and participating organizations and knowledge of the current wider institutional contexts (e.g., goals, organization, leadership, resources, limitations).

We further identified a more specific set of outcomes as relevant within this overall model:

- Increased know-how and know-who
- Persistence and resilience in the face of obstacles, setbacks, and frustration
- Improved ability to successfully guide one's own learning
- Improved ability to complete tasks cooperatively with others
- Production of improved and high-quality products
- Increased sense of agency in a domain
- Increased comfort with participation in a task or domain
- Improved ability to assess the credibility of claims and information sources
- Assisting others to achieve valued learning outcomes
- Widened range of mobilization of learning across tasks, domains, and settings
- Increased emotional maturity and productive use of affective sensibilities
- Continued development and application of learning over several years or more

In the review of the literature that follows this introduction, we have organized these outcomes into four clusters to make the section more readable and manageable, and include examples of some of these outcomes in the exhibit that follows.

Exhibit 2. Clusters of Informal Learning Outcomes by Level, with Examples

Outcomes	Project	Group	Individual
Increasing comfort with and ability to conduct independent inquiry across a widening range of domains, including evaluating sources and contributions.	Colorado Hybrid Project's cultural responsiveness promotes girls' identification with STEM practices.	Families' scientific-sense making at a marine park demonstrated by TOBTOT.	Use of Zydeco for "nomadic inquiry" of content across learning settings.
Improving ability to learn and act collaboratively, including relevant understanding of and support for learning partners.	GIVE project prompts groups of intergenerational museum visitors to engage in inquiry.	Children engage in mutual helping behavior at 5th Dimension sites.	Collaborative problem-solving for success *Lineage* and other MMORPG play.
Improving quality of products, including ability to critically reflect on the quality of one's own and others' productions.	Digital Zoo designed for development of students' engineering epistemic frames	Youth shape one another's programs in Computer Clubhouses.	DUSTY participants develop agentive capacity in creating their digital life-stories.
Increasing range of social resources and networking to achieve goals.	Programming at 5th Dimension sites is sustained through partnerships and scaling of practices.	YouMedia participants work with peers to create products relevant to their shared interests.	WINS participants in museum program draw on resources to help support STEM career paths.

The research projects summarized in the review of the literature were selected for inclusion because they provide examples of methods for documenting and assessing one or more of the above outcome clusters at one or more of the three levels of analysis. In the review, we specify at the beginning of each project summary the outcomes and levels assessed in each project. We also are preparing for the project wiki a version of this exhibit in which we group all the research projects summarized here into a table for each combination of outcome and level.

A Framework of Basic Concepts

The discussions in our expert panels frequently focused on emerging re-conceptualizations of key concepts pertinent to documentation and assessment design for informal learning activities. There was broad consensus across the three expert meetings, but the report's authors assume responsibility for the specific formulations provided here. Some key terms in the individual project studies reviewed in this report will be used differently from how we use them. We will try to make this difference clear in each case, while maintaining our own consistent usage of the following terms otherwise.

Learning: Learning that matters is learning that lasts, and learning that is mobilized across tasks and domains; our notion of learning includes social-emotional-identity development as well as know-how and know-who; it should also include learning by groups and communities or organizations as well as by individuals. (For a discussion of "informal learning," see next section of this report.)

Knowledge: *Knowledge that matters* is knowing how to take the next step, for which declarative knowledge is merely one subsidiary component and greatly over-emphasized in current assessment; *know-that* matters only insofar as it is mobilized as part of know-how; *know-how (cultural capital)* matters for career futures and social policy only when effectively combined with *know-who (social capital)*; the social networking aspects of relevant knowledge are under-emphasized in current assessment.

Know-how and other aspects of knowledge also need to be defined for groups and communities as well as for individuals. *Groups and communities always know more,* collectively, than any individual member, and collective intelligence and problem-solving skills, creativity, and innovation are also generally superior to what individuals are capable of.

Assessment: The production of knowledge useful for individuals, groups, and communities to improve practices toward valued goals; to be distinguished from evaluation.

Evaluation: Judgments made regarding how well goals are being achieved and how valuable the totality of all outcomes is.

Documentation: The collection of information useful for assessment and/or evaluation.

Research: The production of knowledge useful for the design of activities and communities capable of reaching stated goals and with enhanced potential for producing valuable outcomes beyond stated goals.

Assessment, Research, and Evaluation all build on Documentation, but may require different modes and foci of documentation. Assessment aims at improvement; research aims at knowledge useful for future design.

Engagement: Affective involvement in and commitment to an activity, goal, practice, group, or community which enhances the quality and quantity of participation despite obstacles, setbacks, or frustrations; distinguished from Enjoyment.

Enjoyment: The positive feeling accompanying an activity that makes it worth doing for its own sake. Both *Engagement* and *Enjoyment* are important aspects of learning and should be documented in assessment, while recognizing that negative feelings may also play a significant role in engagement and in learning.

Agency: A term that has different meanings, including: actual effectiveness; a disposition toward taking action; a feeling of self-efficacy; or an aspect of one's identity as someone who can produced desired effects. All these meanings are task- and/or role- and domain-specific, and also often group- or community-specific. The notion of agency also extends to what a group or community believes it can accomplish or can actually accomplish.

Outcome: Conventionally, a (sometimes naïve) attribution of a valued condition to some specific cause (e.g., to an intervention). Rarely, however, are valued learning goals the outcome of discrete, identifiable causes. Moreover, posited "outcomes" observed at some moment in time or over a short interval do not necessarily persist or ground further development. They are frequently transitory phenomena, artificially produced by the procedure used to "measure" them. We will instead use the term outcome to refer to *socially and personally valued, on-going processes which emerge in the milieu of some community and its activities.* Note that we regard evidence of learning-in-progress as equally important with evidence of completed or stabilized learning.

Unit of Analysis: What should be the unit in focus in assessment design? We believe that it should be a *system of activities and practices over time*, which includes the actions of individual learners, but also the roles of other participants, including mediating tools and semiotic media as well as local conditions directly relevant to and supportive of (or obstructing) the learning activities. The unit of analysis needs to be extended peripherally to wider contexts that make the learning activity possible institutionally, but with decreasing detail as their relevance to the specifics of learning trajectories decreases. Assessment at the level of individuals, groups, and whole projects are necessarily inter-dependent and assessment design needs to include all three and their relations to one another.

Review of the Literature

Organization of the Research Review

This section of the report provides a detailed review of selected reports on studies and projects across several domains of informal learning.

What is meant in this report by the shorthand expression "informal learning" builds on a fundamental assumption: Learning in its broadest sense takes place in every activity of life insofar as what we do at a later time benefits from earlier experience. Informal learning refers more specifically to learning as an outcome of participation in organized activities in face-to-face or online settings (or both) other than formal instruction, in which a number of the following features are especially relevant or salient:

- Voluntary participation
- Relatively equitable power relations in negotiating goals and means
- Enjoyment of the learning activity for its own sake
- Intense engagement with tasks
- Flexibility of goals and in re-purposing resources
- Unpredictability of some significant learning outcomes
- Improvisation and innovation within and concerning the activity
- Commitment over time in the face of obstacles

This degree of narrowing still leaves the potential scope of "informal learning" too broad for the purposes of this report. Within such a scope, we would include domains such as organized sports activities, learning in practical activities in the home, and leisure activities developed by organizations such as the Boy Scouts— all of which we exclude. The domains that are *included* in this research review are the following:

- Learning in after-school programs and community centers
- Short-term, focused out-of-school activities
- Activities in informal learning institutions (museums, aquariums, zoos, etc.)
- Computer-based and online activities

For each domain or type of setting, the review looks at exemplary studies and projects, identifying particular approaches to assessment, methods of documentation, and techniques of analysis. The types of learning outcomes investigated by the original authors and the levels (individual, group, organizational) referenced by them are identified. After the review, we include summaries of the discussions in the three expert meetings.

Further details on the expert meetings and other resources identified during the project are available on the project wiki: http://documentinglearningworkshop.wikispaces.com/.

Survey of Literature

Afterschool & Community Centers

In this section, we discuss examples of after-school programs run in a variety of locations and with a variety of educational objectives.

After-school programs: General considerations

A common source of social and scientific interest in educational programs conducted during non-school hours arises specifically because the learning activities are *not* occurring in school. Although they vary greatly, the community-based programs described in this section all seek to leverage the fact that after-school settings are more flexible in terms of their schedules and also their social arrangements, allowing children freedom to engage in peer interaction and less hierarchically codified interactions with adults. Because they take place between the home and the school, in a time of day that—since the advent of modern schooling—has given play a privileged position, they are places where having fun is an essential ingredient and where children participate voluntarily.

This shared circumstance offers both the greatest promise and the greatest challenges to realizing the hopes placed on after-school activities for infusing meaningful learning into the lives of children. On the positive side, the organization of after-school settings is overwhelmingly centered on activities that are designed so that children engage voluntarily; if you build it and it's unattractive, the kids simply will not come. One attraction is that after-school programs generally allow children to work in groups with their peers and to choose the roles that they play in the various projects that are offered. Indeed, children and youth often have a voice in the projects that are offered or are allowed to walk away from those they find boring.

In this kind of social environment, children are free to speak with one another, often using languages they feel most comfortable in, and the staff (who vary in age from their late teens to middle age) are freed of the obligation to know all the right answers, allowing them to position themselves in the role of coaches and more experienced peers. As a result, participation is also a learning experience for adult staff, including older-age youth.

Unfortunately, many of the characteristics that offer the greatest potential for after-school activities to promote learning are the same characteristics that offer the greatest challenges for using after-school settings to provide the kind of educational experiences for which they appear to be perfectly suited. To begin with, because after-school activities are not part of the formal school system and it is not legally binding on children to attend, the

funding sources to support such settings are scarce and uncertain, as is consistency of attendance by the school-aged population such settings are designed to serve. For the same reason, few staff earn a living wage from this work, leading to a high rate of staff turnover and a low level of education of the staff with whom children and youth are in close contact. These "structural" factors (Bevan & Michalchik, 2012) go hand in glove with the promising social-organizational characteristics of after-school activities summarized above. If a learner's project goes unfinished, it may well not be taken up again in a later session, creating difficulties for programs with goals such such as promoting science engagement. Often projects do go unfinished, perhaps even un-started, as children, restless after a long day of sitting in enforced quiet in school, are freed up to "have fun" in ways that have nothing to do with pre-planned activities.

For our present purposes, a major challenge of informal, voluntary, peer-oriented, group-organized after-school programs is that they are difficult to evaluate. Although "evidence-based" assessment has meant merely "randomized trial" research designs, as federal regulations have often been interpreted in evaluating formal education, the voluntary nature of informal learning activities precludes random assignment and makes persistence of learners in a trial unreliable. Participants come and go, appear for some sessions but not all, change activity groups as their friendship networks evolve, and leave the activity if their participation is overly controlled. Standardized assessments, moreover, are based on the assumption that valued outcomes of an activity can be known in advance, whereas experience with informal learning programs shows that some of their most valued outcomes emerge unexpectedly.

Each of the programs reviewed here provides useful suggestions for ways to assess informal learning activities. A key lesson they teach is that one size does not fit all. As in the programs themselves, a variety of assessment strategies is the rule, not the exception (e.g. Harvard Family Research Project, 2008; Mahoney, Larson, & Eccles, 2005; Bell et al., 2009; Shaffer & Gee, 2012; National Research Council, 2000).

Examples of After-school Programs

We have chosen to focus on several programs that have achieved sufficient scale to provide a foundation for reaching conclusions that are potentially generalizable across a reasonable range of variation. These programs share the following characteristics:

- A focus on leveraging new technologies for learning after-school
- A university-community partnership model for after-school learning and research
- The use of fieldnotes, interviews, video analysis and other site-specific evidence of learning and development

5th Dimension Program: A broad-based enrichment program

Levels of Analysis: Individual, group, project
Valued Outcomes: Improved literacy and numeracy, agentive participation, long term sustainability
Methods: Cognitive-ethnographic fieldnotes of adult (including college student) participants, videography, quantitative data collected as part of normal practices

The 5th Dimension is an educational activity system that offers school-aged children a specially designed environment in which to explore a variety of off-the-shelf computer games and game-like educational activities during after-school hours. (The account here draws heavily on the monograph by Cole and the Distributed Literacy Consortium, 2006.)

The 5th Dimension involves co-participation of undergraduates enrolled in a practicum course and elementary age school children in a community setting. Such settings are organized so that adults provide for the involvement of children in presumably pro-social, development enhancing, activities in the non-school hours in spaces that can house such activities (e.g., Boys and Girls Clubs, YMCA's, local libraries, and after-school programs in a school, church, community center, etc.)

The 5th Dimension was modeled loosely on the idea of Dungeons and Dragons. The conceptual layout of the 5th Dimension includes a labyrinth that contains a variety of computer and non-computer games that are a part of a make-believe playworld. The 5th Dimension's materials and conventions are organized to achieve the kinds of objectives enumerated above. By intertwining "fun" and "academic" motivations, the activities promote children's engagement in reading and writing. At the same time, this playworld is designed to help participants (both children and undergraduate students) orient to the game's attractions and challenges, to form goals, and to chart progress toward becoming an expert. Other features were designed to provide motivation to write to someone, to look up information in an encyclopedia, to teach someone else what one has learned, and to reflect upon and criticize information.

As a means of balancing play and "work," 5th Dimension sites often have an electronic entity (a wizard/wizardess) who lives in the Internet but writes to (and sometimes chats with) the children and undergraduates online. In the

mythology of the 5th Dimension, the wizard/ess and acts as the participants' patron, provider of games, mediator of disputes, and the source of computer glitches and other misfortunes.

Because it is located in a community institution, 5th Dimension programs require a local "site coordinator" who greets the participants as they arrive and supervises the flow of activity in the room. The site coordinator is trained to recognize and support the pedagogical ideals and curricular practices that mark the 5th Dimension as "different"—a different way for kids to use computers, a different way of playing with other children, and a different way for adults to interact with children.

The presence of university and college students is a major draw for the children. The participating college students are enrolled in a course focused on fieldwork in a community setting where they act as “buddies” for the children. It is in these cross-generational collaborative activities organized around solving puzzles/playing games that make the socio-emotional aspects of the children’s behavior are most clearly visible.

5th Dimension Assessment strategies

Owing to the wide variety of institutional settings where 5th Dimensions have been implemented, an equally wide variety of assessment strategies have been used (see Cole & the Distributed Learning Consortium, 2006, for details). Some of these strategies have involved a focus on outcomes of learning, while others focus on evidence of learning-in-process. Some focus on individual children, some on the success of the activity within its community setting judged by children’s attendance and levels of enthusiasm, some based on the degree of support and participation by the cooperating institutions. The particular mixture of methods depends, as does the activity itself, crucially on its context.

The outcomes-based assessments focused on individuals are most likely to be useful in situations where the local social ecology permits random assignment of children to participation in the 5th Dimension. In such cases external criteria (statewide achievement tests provided by the schools as well as various tailored pre/post test tasks given as a condition for participation) have been implemented. In situations where no proper control group can be included but where there are more children desiring to participate than the facility can handle, various quasi-experimental tasks using 5th Dimension activities have been implemented.

Evaluations at this individual-focused level have included (Cole and the Distributed Learning Consortium, 2006):

1. Computer literacy (paper-and-pencil computer knowledge test, evaluation of memory for computer terminology, hands-on computer-use proficiency merit badges)

2. Mathematical understanding and mathematical problem solving (understanding arithmetic word problems isomorphic to those in the computer games, the use of math problem-solving strategies in the games themselves as evidenced by fieldnotes

3. Reading and linguistic skills (reading comprehension of instructions in novel games, reading notes from and writing notes to the Wizard, engagement in live written chats with the Wizard, and success in a variety of record keeping practices)

Since it is impractical to create control or comparison groups for group-level outcomes, researchers have used in-situ, process-focused evidence of children taking greater control of the conduct of the activities, longitudinal studies of children's progress from beginner to excellent levels of performance within games, data-mining of large samples of fieldnotes to reveal the frequency and sophistication of academic skills such as reading and of social skills such as helping others.

Criteria for assessment of success at the level of the activities within cooperating institutions include whether supervisory personnel who are supported by the community organization participate in the activities at their site, and whether the university/college supports continuing practicum classes that supply supervised undergraduates as "more capable peers" or "buddies" to work/play with the children and write detailed fieldnotes. To attain its long-term goal of a sustainable new cultural practice, the cooperation of the implementing institutions in obtaining the additional funding needed for the transaction costs of the collaboration provides a strict criterion for this objective.

Across dozens of 5th Dimension programs, each using a mixture of appropriate evaluations of the kind sketched above, it has been shown that the program is effective at the individual level. Children like the activities and participate, they are generally agentive, argumentative, and engaged in observable and documentable ways. In a great variety, but not all, of the social contexts where the innovation has been tried adults approve of the activities and make their children available. College students show marked improvement in understanding how to be effective supporters of children's and their own learning. At the group level there is evidence of improved group collaboration and task success. (Cole & Distributed Literacy Consortium, 2006; Downing-Wilson, 2006/2007).

One of the instructive features of the 5th Dimension project is the extent to which it has afforded assessment of learning at the level of the program as a whole in its socio-ecological contexts. No 5th Dimension is a static, unchanging system of activity; instead, each changes in relation to itself and its institutional settings, which are themselves changing. For example, at one implementation of the 5th Dimension at a Boys and Girls Club over a 16-year period (Downing-Wilson, Lecusay, Rosero, Cole, 2012), the careful documentation process reveals a dynamically changing relation between the "5th Dimension itself" and its host institution. Initially the 5th Dimension is more or less self-contained in an accessible room. Then, over time, in response to changes in Club policies and resources, the practices of the 5th Dimension gradually start to infuse all areas of Club activity, engaging a wider range of children and engaging Club personnel more deeply in the program and enlarging the role of the 5th Dimension in its setting.

Other 5th Dimension programs show different patterns of learning at the project and organizational level as they adapt to rapid changes in computer technology, turnover in personnel, changing funding streams from a variety of sources, and the changing priorities of the collaborating institutions. Unsurprisingly, there are failures to adapt as

well as successes. Many 5th Dimension projects run successfully for two, three, or more years, only to fall apart owing to changes in institutional priorities or the departure of key personnel. Others manage to adapt to the changed conditions and have continued for many more years. At the time of this writing, some such systems have been in existence for more than 15 years permitting analysis of the institutional and social-ecological conditions for sustainability.

Learning to Program in a Computer Clubhouse

Level: Individual, group, project
Valued Outcomes: Collaboration, quality of products, social resources
Methods: Case study, Oakes' model of school reform, portfolios of activities

The Computer Clubhouse Network is another after-school program similar in organization and goals to the 5th Dimension program. Started in 1993 through a collaboration between MIT's Media Lab and Boston's Computer Museum, the network currently consists of over one hundred after-school learning centers dedicated to providing primarily high-end digital media instruction to help elementary-school through high-school age youth develop the expertise to express themselves fluently through new technologies (Resnick, Rusk, Cooke, 1998). To create and sustain an environment in which this communicative fluency is developed, Computer Clubhouses are designed to cultivate an "emergent community" that promotes opportunities for youth to engage in design projects that are driven by their own interests (Kafai, Peppler, & Chapman, 2009).

Of the several innovative and technology-centered learning activities that have been developed and field-tested in Computer Clubhouses, the *Scratch* programming activity has generated particular attention. Scratch is a programming environment designed to help novices learn programming through easy desktop manipulation of digital images and sound files. Rather than type code, learners program by dragging and dropping "blocks" of preset programming commands in a jigsaw puzzle-like fashion, creating "stacks" of blocks. The commands control the actions of movable "sprites," each of which contains its own set of images, sounds, variables, and scripts (see Exhibit 3 for an image of the Scratch interface).

Exhibit 3. Screen shot of the Scratch Programming interface.

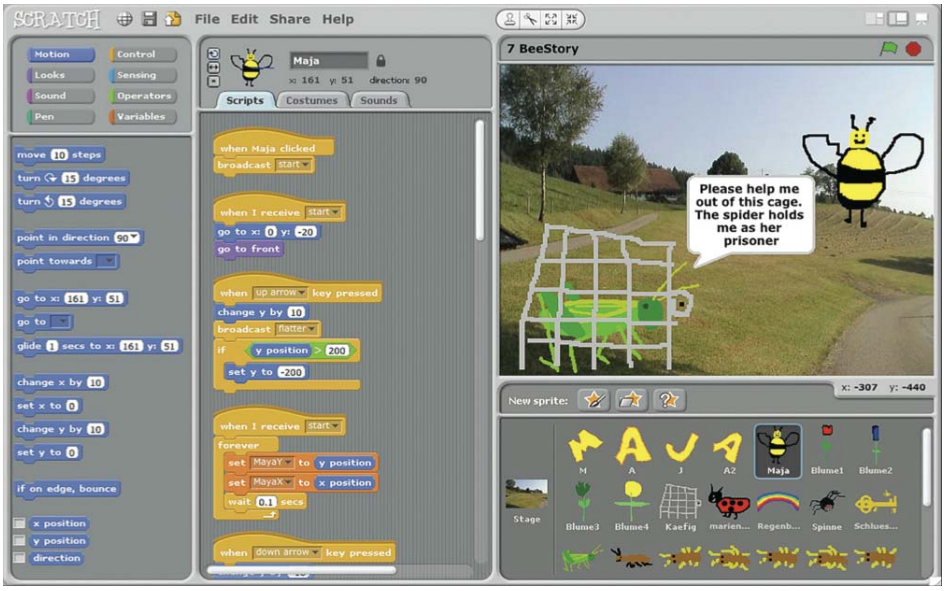


Image from Resnick et al. (2009).

Kafai, Peppler, and Chiu (2007) examined the organization of activities at a Computer Clubhouse before and after the introduction of Scratch as one of the activities in order to understand how after-school learning environments might be "seeded" with a computer-programming activity. Kafai and her colleagues drew on Oakes (1992) model of school reform. According to Oakes, efforts to reform educational activities in a manner that insures that they are equitable need to go beyond *technical* aspects for implementing these activities. It is also necessary to encourage changes in the *normative* and *political* dimensions of the activity—that is, its longstanding norms, values, and institutional support within larger community.

Kafai and her colleagues documented and analyzed the organization of activities at two Computer Clubhouse sites in Southern California. At the technical level, they were interested in observing how, and to what degree, Scratch would be adopted as an activity given specific conditions of implementation. These conditions included the incorporation of new activities like workshops and gallery presentations and the participation of undergraduates who acted as mentors but who knew little to nothing about programming. Paralleling principles of the 5th Dimension model, the inclusion of inexperienced programmers as mentors was seen as contributing to the occasioning of opportunities for mentees to teach the novice mentors about programming in Scratch (Kafai et al., 2007).

Normatively, the focus was on how participants (including youth, mentors, and parents) interpreted the meaning and value of computer programming (e.g. what did participants consider prototypical programming projects?) and how they interpreted the value of their own programming abilities. With respect to the political dimension of the analysis, Kafai and her colleagues report the way in which a formal partnership developed between the University

and the Computer Clubhouse's host organization. These changes were crucial, they note, in gaining the support needed to create the infrastructure necessary for the Computer Clubhouse's goal of achieving technological fluency.

The primary documentation data at the individual and group levels were researcher and mentor-produced fieldnotes and sample products of participants' efforts. To address technical and normative questions about how Scratch was incorporated in the Clubhouse and the role of the mentors in this process, the fieldnotes were coded for different kinds of activities *(design activities, web activities,* and *personal activities, e.g. socializing).* The products produced at the Clubhouse including projects of all kinds (including *Scratch* programs) were analyzed based on project types such as animations, games, stories, and graphics (Kafai, Peppler, Alavez, & Ruvalcaba, 2006). In addition, group interviews of Clubhouse members and undergraduate mentors were conducted in order to capture how the young participants understood the development of their own programming skills.

From the analysis of fieldnotes, projects, and interviews, *portfolios of activities* were generated to show changes in the kind and number of activities the youth engaged in before and after the introduction of Scratch. The identification of periods of increased production guided more focused analyses of the kind and distribution of projects. These steps, in turn, allowed Kafai and her colleagues to identify and conduct case studies of Clubhouse members who created projects, the structure and content of which appeared to influence the production of Scratch projects by other Clubhouse members. (When we consider computer programming as an activity in itself in a later section of this review, we will return to examine how researchers also tracked the quantity and distribution of programming commands in Scratch projects over a period of time).

Digital Storytelling & Media Production Programs

In this section we examine two additional projects, each of which seeks to create environments in which youth gain exposure to new, academically relevant activities and technologies in ways that encourage the incorporation of the learners' interests. Both examples focus on assessing learning through careful examination of the process and products of youth efforts to produce digital media. The first case, a digital storytelling after-school program, illustrates ways in which extensive ethnographic data, including participant produced media, can be analyzed for insights into the development and enactment of learner's authorial agency. The second case presents an example of how data visualizations can be used to assess learning in participants' uses of online social media platforms in a technology-driven after-school program.

The Digital Underground StoryTelling Youth Project (DUSTY).

Level: Individual, learning ecologies
Valued Outcomes: Independent inquiry, collaboration, social resources, identity formation
Methods: Case Study, thematic analysis, narrative analysis, performative moments

The Digital Underground Story Telling Youth (DUSTY) Project is an adult and youth multimedia literacy program run out of a community technology center located in the Bay Area. DUSTY differs from traditional, academically-

oriented after school programs in its emphasis on the centrality of identity formation and meaning making in learning, particularly the role that semiotic systems in addition to language play as resources for "embodying and enacting a sense of self in relation to others," (Hull & James, 2007). To this end, the creation of digital stories at DUSTY is aimed at positioning participants to use narrative reconstruction to reflect on their experiences and to be active agents in articulating their own aspirations.

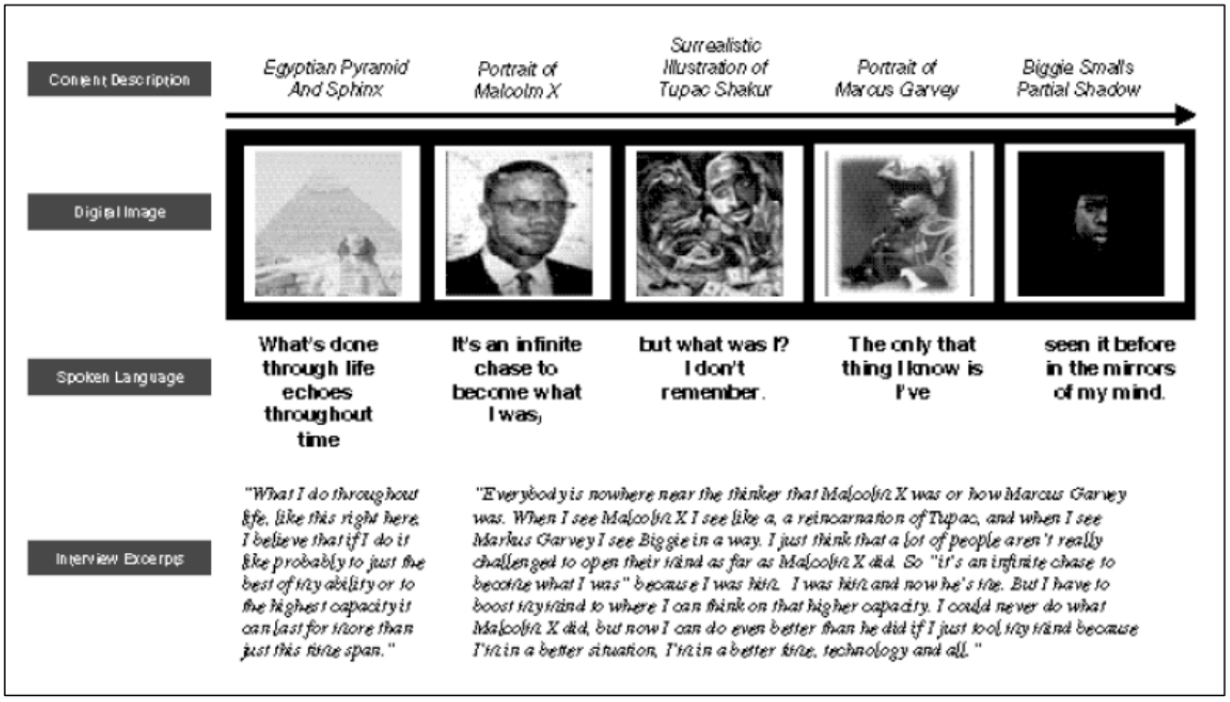
This agency and identity-centered framework for theorizing and assessing learning is described in detail by Hull and Katz (2006). The authors emphasize the importance of the various social contexts of learning (e.g. the DUSTY program, the school) and their contributions to a learner's digital story production. They argue that people can use their personal repertoire of tools and other resources to develop agentive selves by creating "multi-media autobiographical narratives about self, family, community, and society … to articulate pivotal moments in their lives and to assume agentive stances toward their present identities, circumstances and future".

Hull and Katz report longitudinal case studies of youth participants using as documentation: fieldnotes of participant activities across multiple settings, written by researchers and undergraduate mentors; interview data to characterize participant histories inside and outside of DUSTY; and participant produced media including story boards, scripts, and digital stories.

To demonstrate how two such participants developed authorial agency through their creation of digital stories, Hull and Katz drew on a variety of concepts, originally proposed by scholars such as Bruner, Urciuoli, Bauman, and Briggs, to identify displays of agency both in these youths' actions (captured through fieldnotes and interview transcripts) and in their digital stories. They employed Bruner's idea of narrative "turning points," when people "report sharp change in their lives and demonstrate accompanying dramatic changes in their representations of self" (Hull & Katz, 2006, p.45). From Urciuoli (1995), they adopted the idea of performative moments, which include "any activity that coordinates action to create a unity from many selves" (p.202). From Bauman and Briggs (1990), they adopted an "agent-centered" view of verbal performance in which agency is exercised by a person through the decontextualization and recontextualization of verbal texts, and applied it to the multiple modalities of digital stories (see Exhibit 4).

The case study subjects not only mastered the technical skills needed to create digital stories, they also became more sensitive to the genres and poetic aspects of language, and they grew increasingly adept at combining multiple media (text, sound, and images) to create personally relevant narratives.

Exhibit 4. Excerpt of graphic representation of one of Randy's digital stories



The representation was used as a tool to identify and analyze patterns relevant to the focus of analysis in Randy's case, authorial agency. The representation juxtaposes screen shots from the digital story, text of the corresponding voice over, and excerpts of relevant comments from interviews. Borrowed from Hull & Katz (2006), p. 51.

Data Visualizations to Assess Learning in Online Social Learning Networks: iRemix and YouMedia.

Levels: Individual, group
Valued Outcomes Independent inquiry, collaboration, quality of products
Methods: Social learning network analysis, timeline structures for data visualizations

YouMedia is another example of a program dedicated to providing youth with the space, resources, and semi-structured opportunities to use high-end digital media to explore new modes of self-expression. Housed in a centrally-located urban library, YouMedia is a collaboration between the Chicago Public Library and the Digital Youth Network, a Chicago-based digital literacy project that provides linked in-school and after-school media classes (http://www.digitalyouthnetwork.org/). At YouMedia teens from around all parts of Chicago have access to a rich array of technological and social resources, including laptops, smart boards, video cameras, and an audio recording studio.

Based on similar theoretical principles as DUSTY (see above), activities at YouMedia are organized to encourage teens to identify topics of interest and to explore telling stories about these topics through multiple media. Mentors from the Digital Youth Network visit YouMedia to share their media production expertise with the teens, who work on digital activities in music, design, photography, blogging, and video production.

Teens who wish to participate in YouMedia activities must register as members. As part of this process, the teens are invited to join a cloud-based, online social network, YouMedia Online, restricted to and accessible only by YouMedia members.[1] Users of YouMedia Online can do things common on social networking sites, including managing their profile pages, posting comments and media, and joining groups.

However, YouMedia online differs from other social networking sites in that it is designed to support both structured and self-directed learning in ways that connect to the culture and expectations of the YouMedia community (Austin, Ehrlich, Pucket, & Singleton, 2011). For example, one of the ways that learning is promoted is through the use of a virtual currency called "Remix dollars." These dollars are awarded for, among other things, posting original content and providing feedback to others by posting comments on their projects. This currency can then be redeemed for such items as gift cards, USB drives, MP3 players.

As a repository of vast amounts of information about the teens' YouMedia-related activities, YouMedia Online offers a rich source of data for exploring questions about learning and the development of expertise, interests, and social capital (Nacu, Pinkard, Larson, Schmidt, 2012). Nichole Pinkard, who founded the Digital Youth Network, and her colleagues have taken the first steps to exploring the mechanics, ethics, and potential insights of mining data from social learning networks like YouMedia Online (Nacu et al., 2012). They have examined ways of adapting and applying for the YouMedia Online platform the tools, metrics, and analytic frameworks typically used in the private sector for studying the use of online social networks.

The questions that are driving Pinkard et al.'s research include:

- How are youth participating and interacting with others?
- How do we know that such interactions lead to learning?
- How do interactions with peers and mentors play a role in learning?
- How do actions such as viewing media and user profiles relate to production-oriented actions such as posting comments, critiques, rating media, and posting original work?
- What patterns of passive and active participation are visible?

Additionally, Pinkard and her colleagues have identified features of social learning networks that can be studied to address questions concerning the kinds of learning that unfold for users of these networks (see Exhibit 5).

---

[1] The platform for the online social network is *Remix World* (http://remixlearning.com/). Developed by Nichole Pinkard and her colleagues, the platform was designed specifically for use by educational institutions—schools, to museums, to after-school centers—and is customizable to fit the particular needs of each institution

Exhibit 5. Features of online social learning networks (SNL) and their corresponding forms of data

Participation	Consumption	Contribution	Production
Access and membership	Views of specific items	Content posted to the network	Posting original media
Impact	**Relationships**	**Expertise**	**Interests**
Evidence of impact by individuals	Social ties, interactions, and group affiliations	Evidence of skill development	Emergence and development of interests

From Nacu et. al. (2012)

In order to make sense of the large amounts of documentation generated from logs of YouMedia Online activity, Pinkard and her colleagues developed operational definitions of actions and concepts they deemed relevant for addressing their research questions (see Exhibit 6). Using these definitions as tools for coding the data, they experimented with a number of techniques to visualize the results. For example, to explore questions concerning the role that peer-peer and peer-mentor relationships play in facilitating learning, social network maps were generated to chart these relationships and their strengths at specific points in the history of the program and over time. Another technique creates timeline structures that show the kinds of activities individual participants have engaged in over extended periods of time and also how often and for what period of time they have been involved.

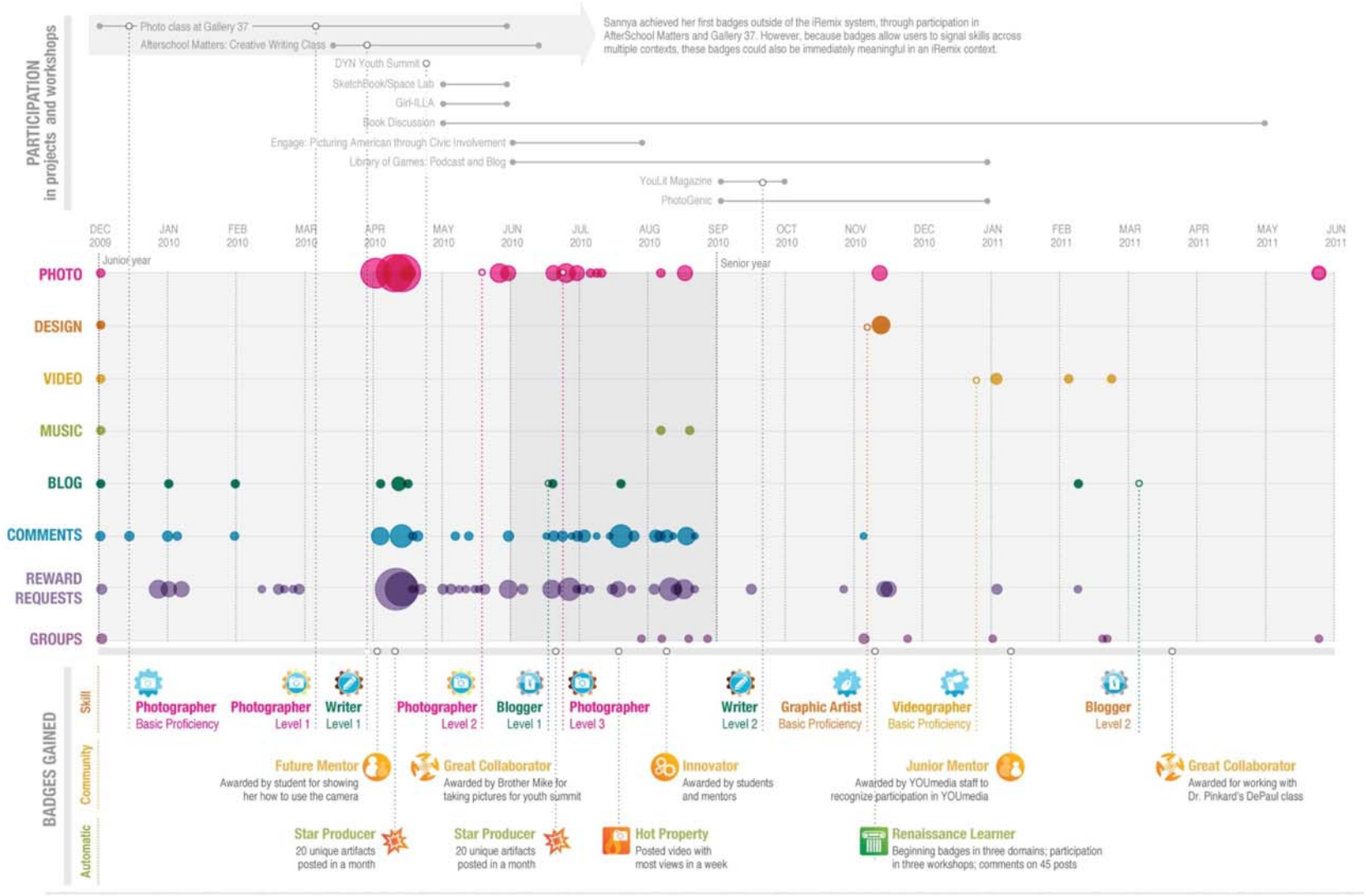
Exhibit 7 shows an example of a timeline structure for a YouMedia participant's online activities during her junior and senior years in high school. The timeline covers a period of 18 months and is divided into three rows. The top row identifies YouMedia projects and workshops that the student participated in, indicating for each the beginning and endpoints of participation. The middle row specifies activities (named on the far left) that the student engaged in and that could be extrapolated from the students' activity on the YouMedia Online network (e.g. editing, uploading and commenting on videos and photos, posting blog entries, joining groups), with the size of the circle indicating the number of digital products uploaded. The bottom row describes badges that could potentially be awarded to the student for engaging in activities in a manner deemed successful or valuable by the local community. (For more discussion of plans to adapt a badges initiative in YouMedia see the Future Directions section below.)

Exhibit 6. Definitions of actions and concepts Nacu et al (2012) determined relevant for addressing their research questions

SLN features	Kinds of data
Structures to support informal and formal interaction among users	Social ties
Access to teachers/mentors/expert	Interaction patterns among peers/ teachers/mentors
Asynchronous, online communication	Access and participation patterns
Creating, sharing, and discussing multimedia content	Engagement with content Connecting use patterns with patterns of learning and development
Ability to structure learning activities and projects	Engagement and impact of specific learning resources

Figure borrowed from Nacu et. al. (2012)

Exhibit 7. From Nacu et. al. (2012)



STEM-focused Community Based Programs

The projects discussed in this section differ from those above in that they were designed with specific content-area learning goals. In each case the content area was in the domain of science, technology, engineering, mathematics and computer science (STEM).

Emergent Learning in a Community Garden Program

Levels: Individual
Valued Outcomes: Independent Inquiry, Social Resources
Methods: Case study, semantic domain analysis, discourse analysis

Jrene Rahm (2002) conducted an ethnographic study of an 8-week summer community youth program designed to teach plant science and entrepreneurship, as well as to develop participant's team-work skills. Based at a community garden site in an ethnically diverse low-income, inner city neighborhood, the program was structured so that participants rotated through 2-week cycles in four kinds of activities: nurturing (soil preparation, planting, watering), harvesting (identification, harvesting, and preparation of harvestable crops), marketing (contacting and selling to potential buyers, organizing delivery), and special projects (community outreach, artwork, tree planting).

The aim of the study was to examine how opportunities to learn science emerged in the everyday interactions among the program participants. To this end, Rahm focused her research on the language and discursive practices that the participants used to make sense of program-related activities.

As a participant observer, Rahm simultaneously videotaped, kept fieldnotes, and worked alongside a team of six 11 - 14 year old program participants as they engaged in the nurturing and harvesting activities. Following the data collection phase, transcripts were made of the videos and integrated with corresponding fieldnotes. Rahm subjected these materials to domain analyses (Spradley, 1980) in order to develop taxonomies of the kinds of learning opportunities supported by the program. These analyses revealed some of the discursive strategies that participants used to make meaning of the gardening activities (e.g. analogies, elaborations, questions, and comments). Rahm found that question strategies could be further subdivided into information questions, knowledge integration questions, and inquiry. In 95% of the cases studied, inquiry and knowledge integration questions were precursors to learning opportunities.

Rahm also explored the connections that the youth participants made between their experiences in the program and their experiences with science in school. She conducted pre/post- semi-structured interviews of the six youth participants that she shadowed and of select youth participants from the remaining teams of youth participants in the program. The interviews focused on the participants' notions of science, and on their perspectives regarding learning science in the program and at school. Post-interviews were also conducted with all of the adult program participants in order to document their perspectives on the learning that takes place in the program. The interview data revealed that both the youth and adult participants saw the activities of the program as providing authentic opportunities for learning science that were markedly different from those offered in the classroom.

Colorado Hybrid School-Community STEM Project

Level: Group, Project
Valued Outcomes: Independent inquiry, collaboration
Methods: Design experiment, semantic domain analysis, vignette analysis

Eisenhart and Edwards (2004) implemented and studied an after-school program designed to promote participation in computer technology and science activities by urban, African- American middle school girls. This university-community collaboration involved university researchers, representatives of community institutions, and local residents. The program also involved the iterative design and implementation of culturally sensitive science and technology-focused classes in the after-school hours. Of the cases discussed thus far, this is the only one in which a deliberate effort was made to create a learning activity that met state curriculum content standards while at the same time incorporating youth participants' own goals and interests.

This research was organized to examine two issues: (1) the extent to which the participating girls' gender and ethnic identities were represented in the technology-driven activities of the program and (2) the implications of this for understanding how to successfully engage the girls. The research asked: What were particularly good examples of occasions when the girls seemed motivated to learn more science or technology, and what did these occasions suggest about the conditions for successful work with the girls?

To address these issues, Eisenhart and Edwards followed a group of six African-American girls (12-15 years old) as they participated in science and technology activities in the after-school program, over the course of one academic year. Documentation of these activities included researcher- and instructor-produced fieldnotes and journals, audio and video transcripts of each session, audio transcripts of student interviews, the digital and concrete products (e.g. T-shirt designs) created by the students using software provided to them, and computer logs of the students activities using these software applications.

To address the degree to which these girls' identities were reflected in the science and technology activities, Eisenhart and Edwards drew on Spradley's (1979, 1980) semantic domain analysis. The basic procedure involves identifying and coding semantic domains (participants' units of meaning) and constituents (the items that constitute the domains) through the application of nine universal semantic relationships to the data texts. Eisenhart and Edwards examined three domains that reveal how the girls marked gender identity, ethnic identity, and technological expertise: (1) statements and actions regarding interest in computers and technology; (2) actions regarding gender identity; and (3) statements and actions regarding African-American identity. By reviewing the contents of each domain and then examining the overlap between domains 1 and 2 and domains 1 and 3, they obtained a general picture of how markers of gender and ethnicity intersect with activities involving technology.

Vignette analysis (Maxwell, 1996) was used to investigate the question about occasions and related conditions in which the girls demonstrated motivation to engage in science and technology learning. In vignette analysis, salient events that bear on the research question are excerpted from the data texts, simplified, related to each

other, and retold in the form of short stories called “vignettes” (Van Maanen 1988). The object is to develop a description of the context of the activity that links the data to a “coherent whole.”

Eisenhart and Edwards’ analyses revealed the after-school program to be a hybrid space between the formal space of the school classroom and the informal space of students’ social lives outside of school, where participants are free to move outside the roles expected of them and where participation can be more symmetrical (Gutiérrez, Rymes & Larson, 1995). This created opportunities for the girls to make connections between science and their own lives and values, connections that Eisenhart and Edwards found helped sustain the girls’ engagement. They also found that the absence of such connections made it difficult for the girls to sustain engagement.

Interpreting students’ metacognition in an amusement park physics program

Levels: Individual, group
Valued Outcomes: Collaboration, independent inquiry
Methods: Interpretive case study, semi-structured interviews, pre/post assessments.

Anderson and Nashon (2007) studied an informal learning program that was organized to support in-school classroom work. The program took high school juniors and seniors who were enrolled in physics classes to an amusement park where they were divided into teams to assess the kinematics of three rides using accelerometers, stopwatches, and protractors. This amusement park activity was combined with a classroom assignment in which the students, working in their original teams, were given prompts to encourage them to develop explanations, arguments, and models of the physics principles in the rides.

Grounded in an interpretive case study approach (Gallagher & Tobin, 1991; Merriam, 1998; Stake, 1995), the study aimed to assess the students’ metacognitive abilities (individually and as groups) and how these abilities influenced the development of the students’ conceptual understandings of kinematics.

While the authors describe metacognition and learning as idiosyncratic and dynamic processes, this research is centered on the psychological construct of metacognition, measured through a 53-item baseline questionnaire the researchers developed that allowed them to profile each student according to constituent components of metacognition: awareness, control, evaluation, planning, monitoring, and self-efficacy. These baseline metacognitive levels became an important lens through which the researchers interpreted individual students’ discursive interactions in subsequent activities, which in turn contributed to the validation of the baseline instrument. After taking the baseline survey, the 11th and 12th graders in the study visited and conducted their experiments in small groups at the amusement park. During the visit, the research team developed fieldnotes, video recordings of several small groups’ interactions, and individual audio recordings of conversations of the students participating in the in-depth case studies. Case study students were given copies of their personal conversations and asked to listen to them in advance of group interviews about their metacognition and learning. Students participated in one group interview after the fieldtrip, some in-class activities related to the physics

explored on the fieldtrip, and a second group interview after the in-class activities. Audio and video recordings were made of all these events.

In their analysis, Anderson and Nashon interpret students' discussions in learning activities and interviews through the lens of their particular metacognitive profiles. The researchers illustrate how these profiles interact with students' circumstances, dynamically influencing social processes within their small peer groups and concurrently influencing their individual and collective formation of physics concepts.

Museum-based Programs and Projects

In this section we describe several museum-based, informal learning projects. Informal science institutions such as science and technology museums, aquaria, and zoos have long offered both specific educational programs for children and youth as well as educationally oriented general exhibitions for the public of all ages. Research studies on learning in such institutions have focused on individual and group (especially family) learning and engagement with exhibits as well as learning in more specialized programs. Because these settings do not especially afford opportunities for testing visitors' knowledge before and after, studies have often developed other means for assessing learning and other criteria for valued outcomes.

GIVE: Facilitating family group inquiry at science museum exhibits

Levels: Individual, group, project
Valued Outcomes: independent inquiry, collaboration
Methods: Experimental design, video analysis

The GIVE research project was launched at the San Francisco Exploratorium, a leading informal science learning institution (Gutwill & Allen 2010). The project aimed to answer the following questions:

- Can intergenerational groups of museum visitors, such as families, be coached by museum staff to learn a set of inquiry skills that they can use on their own at novel exhibits or even experiences beyond their visit?
- How does such an intervention affect visitors' inquiry behaviors at a novel exhibition it support them to explore the exhibit more deeply?
- What properties of the staff-mediated intervention seem central to its design, and what is the evidence for them?

To probe these questions, the GIVE team developed two inquiry games based on the following learning science-derived design principles: building on prior knowledge; teaching via modeling, scaffolding and fading; identifying skills explicitly; supporting metacognition; supporting collaboration; making the activity intrinsically motivating; minimizing the cognitive load; supporting visitors' learning agendas; and supporting individual spontaneity (the last four of these are specific to informal learning design). Because of constraints in the museum setting, the games the researchers developed targeted only the following two among the full range of inquiry skills:

- *Proposing Actions*: asking a question or making a plan at the beginning of an investigation
- *Interpreting Results*: making observations, interpretations, or explanations during or after an investigation

The researchers chose these skills because they were simple enough to learn quickly and were not ones commonly observed already in museum visitors' behaviors. The researchers additionally viewed these two skills as possible gateways to other inquiry skills such as questioning, predicting, analyzing, and explaining. Along with *experimentation*, an activity that visitors engage in spontaneously at interactive exhibits, *proposing actions* and *interpreting results* form a complete, if simplified, version of the inquiry cycle as it is typically presented in school curricula and characterized by science researchers.

After iteratively refining the games, the GIVE team conducted an experimental study comparing the two game conditions with two control conditions. They studied 50 families in each condition, 200 families total, to determine if the games promoted inquiry behavior at exhibits. Families in the treatment groups learned how to play the games with supports at two or more exhibits and then, as a post-test, were asked to use a final exhibit playing the game they had learned. Afterwards, the research team conducted interviews with one adult and one child per family selected at random. Researchers also collected demographic data regarding each participant.

The GIVE team video-recorded all families at their first and last exhibits and coded participants' behavior according to the following codes: engagement, proposing actions, interpreting results, collaborative explanations, and coherent investigations. They also coded the data for the correctness of the science content. Results showed that the inquiry games succeeded in improving participants' inquiry. The more structured and collaborative of the two games had the strongest effect, leading to longer engagements with the post-test exhibit, more abstract interpretations, and more collaborative and coherent inquiry investigations than controls. Based on their qualitative analyses, researchers attributed the greater success of this game to its inclusion of all family members in collaboration and its supporting their explicit articulation of their interpretations of the exhibits.

TOBTOT: Assessing museum learner talk over time

Levels: Individual, group
Valued Outcomes: Collaboration, independent inquiry
Methods: Discourse analysis, thematic analysis

This study (Ash et al., 2007) was designed to explore dialogically-based methodological approaches, in particular the authors' Tool for Observing Biological Talk Over Time (TOBTOT). The tool was used to analyze scientific sense-making in biological talk in out-of-school settings and track the language people actually use in discussing life science themes over time.

The researchers' methodological concerns stem from several theoretical traditions, exemplified by the work of Vygotsky, Halliday, Wertsch, Bakhtin, Lemke, and others that take a sociocultural view of conversation as both structured and dialogically emergent. Their applications of the TOBTOT framework focused on testing ways to

track thematic patterns in order to analyze, in both quantitative and qualitative terms, the development of scientific content in dialogic interaction. Ash et al. focus on the methodological tensions entailed by working along the dimensions that encompass content and process, everyday language and canonical scientific discourse, macro- and micro-levels of analysis, and qualitative and quantitative representations of data.

During a 3-year period, the research team audio- and video-recorded 20 Spanish-speaking and English-speaking family groups as they interacted for 20 to 80 minutes at exhibits in a marine biology center in northern California. The families were recruited from a local Head Start center and each included at least one parent and two pre-school to elementary-aged children. The researchers also conducted interviews with family members after the data collection sessions at exhibits, asking for reflections on selected video clips of their naturalistic interactions as a "member check." The analytic tool used, the TOBTOT, consists of three superordinate thematic categories for coding dialogue related to biological phenomena: Staying Alive, Characterizing, and Ecological Interdependence. Staying Alive roughly corresponds to the concepts and discourse traditionally associated with the discipline of biology. Characterizing is the TOBTOT category for talk related to labeling or locating an animal. Ecological Interdependence categorizes talk thematically related to habitats, human impacts, and, less typically for life science disciplines, feelings and aesthetics (e.g., expressions of amazement or personal attraction to an organism).

On a macro-analytic level, TOBTOT's categorization scheme allowed researchers to illustrate and provide graphic representations of how families discursively engage with abstract biological content differently from one another, as well as across exhibits, and over time. The research team reported TOBTOT-based evidence suggestive of several interesting findings and noted their implications:

- The naturalistic family dialogue centered on Characterizing and Ecological Interdependence more than Staying Alive talk, emphasizing the importance in everyday talk—and science learning in informal settings—of becoming familiar with and feeling for living things.
- Characterizing and Ecological Interdependence tended to occur prior to Staying Alive talk at an exhibit, suggesting that "naming, using prior knowledge, and making personal connections must occur before the formal science can begin".
- Particular themes occurred as leitmotifs that changed over time, illustrating that "questions (e.g., about feeding) can permeate time and context, often acting as a central core of dialogic negotiation."
- The touch tank generated the most family talk, prompted the researchers "to investigate the special role of living things as mediational means in dialogic activity."

The researchers also illustrated the ways in which the TOBTOT could be used to analyze and document differences between families in their patterns of discursive interaction at exhibits.

The TOBTOT proved useful in helping researchers demonstrate science content in everyday language, largely because it accommodates talk in the Characterizing and Ecological Interdependence categories. Ash et al. point out that "the general observations these families made were predicated on scientific principles" that are part of "cultural scripts based upon common scientific understandings." However, they also note the difficulties they had in determining whether some instances of everyday talk were scientific or not. They further noted that although their data indicated that the "families used multiple resources to talk and act in new ways, and the use of these resources were dynamic, social, and discontinuous, not linear, direct or clear," the TOBTOT was not able to capture the nuances of the families' dialogic interactions and the scientific insights they supported. The authors

also reported that use of the TOBTOT did not resolve tensions related to particularistic treatment of conversational data or readily facilitate conversational segmenting. Therefore, the research team struggled with traditional standards of reliability and the quantitative representation of qualitative data. The researchers conclude that the TOBTOT is a powerful tool for creating abstract interpretations and representations of actual lived experience in order to answer relatively simple questions. The TOBTOT did not help answer the hardest questions, especially regarding the cultural resources people use to make sense of and attach value to natural phenomena in their world.

The Zydeco Tool: Technologies for guiding inquiry, interfacing museum and classroom, and facilitating assessment

Levels: Individual, group
Outcomes: Independent inquiry, collaboration
Methods: Analysis of computer system logs, ethnographic fieldnotes, interviews

Chris Quintana and his colleagues at the University of Michigan's Center for Highly Interactive Classrooms, Curricula, and Computing in Education have developed projects that combine tools and corresponding curricula for helping students engage in scientific inquiry in activities that combine museum and classroom activities (Quintana et al, 2010).

The Zydeco project is one example of this work. It consists of three components: (1) a website where students develop and access study questions, hypotheses, and data, (2) a data collection/annotation component implemented as an iPhone/iPod app, and (3) an explanation construction component implemented as an iPad app "where students view and use the data they and their peers have collected to construct a scientific explanation addressing the questions they are exploring (Kuhn et al. 2012). The system was designed for *nomadic inquiry*: structured inquiry across a variety of settings, including classrooms, museums, the outdoors, and in homes (Hsi, 2003). The Zydeco system is designed to help learners easily transition from one setting to the next including transitions between virtual (handheld devices) and physical (e.g. museum exhibits) contexts.

Kuhn et al. (2012) examined the role that annotations played in how students searched and evaluated data, and how students identify and select evidence that they use in their explanations. They followed a group of middle school students, grouped into pairs, who used the Zydeco system to collect data at a natural history museum for use in a classroom-based project. The latter project involved the students using the explanation construction component to examine their data along with all of the data collected by other students during the museum visit, and using it to formulate their own explanations. Kuhn et al. (2012) found that the annotations made by the students did support inquiry activities, including data interpretation, identification, and search.

The documentary and assessment methods used for this study were a combination of interview data, fieldnotes, and usage logs taken from the Zydeco system. Six pairs of the 54 students followed were given approximately five-minute, daily in-process interviews using a semi-structured interview protocol while they used the a tablet computer to do the explanation exercise. The log data was studied together with researcher generated fieldnotes

of the classroom activity. This included data collected by the students in the museum, which allowed the researchers to study the characteristics of the data including the type of data (e.g. photos, audio notes, photo-audio note pairs) and the accuracy of the annotation titles, tags, and audio notes. A rubric was used to grade the students' final explanations.

WINS: Educational and career trajectories of participants in a Science Museum-based after-school program

Levels: Individual, group
Valued Outcomes: Social resources
Methods: Longitudinal, survey and interview-based study

Fadigan and Hammrich (2004) conducted a study to describe the educational and career trajectories of high school women who participated in a structured museum-based natural science enrichment program know as WINS (Women in Natural Sciences). Developed for academically talented women from urban, low-income, single-parent families, the program's aim was to give participants "the information, encouragement, and confidence they need to consider pursuing careers in the natural sciences, to make informed decisions, and to shape their own futures." In addition to providing participants and their families with free access and transportation to the museum, the year-long program put the participants through summer classes on environmental science, fieldtrips to local parks, and extended stays at an environmental education center. During the academic year, students met at the museum once a week, attend monthly fieldtrips to other science education institutions, and meet scientists.

In order to describe the overall character of the educational and career trajectories for participants in the program, Fadigan and Hammrich selected a sample of former participants, examined the application materials submitted by these participants, and from these materials identified the educational and career trajectories that these participants imagined for themselves. Surveys designed to capture details about the participants' actual educational and career trajectories (high school to the present) were mailed to students in this sample. Following analyses of the surveys a sub-sample of representative participants were invited to come in for interviews. The interview questions were developed to assess how and why the former participants pursed the educational and career paths that they did, and to identify what if any impact WINS had on the development of these paths. The design of the study and study instruments was based on Eccles (1994) model of achievement related choices, which posits that an individual's choices about what career paths to pursue are constrained by the knowledge they possess about possible options; perceptions of gender/role positioning; and perceptions of the value and importance of the options.

The type of findings regarding valued outcomes that this approach permitted included:

- 93% of participants enrolled in a college program following high school completion
- Careers in medical or health-related fields followed by careers in SMET emerged as the highest ranking career paths, with about 20% of respondents pursuing careers in each of these areas.

The majority of participants perceived having staff to talk to, the job skills learned, and having the museum as a safe place to go as WINS elements that influenced their educational and career decisions.

Computer-based and Online Activities

Computers make it possible to keep track of what a person does when they use the machine. Whether based in a computer simulation, an online virtual multi-player virtual world, or a simple text editor, it is now possible to comb through log files or through lines of code to generate records of every action a person takes when using these kinds of software. This information, when interpreted using the relevant conceptual frameworks, can reveal something about what and how a person is or is not learning. A key advantage of this approach is that it permits learning assessment without disrupting the flow of learning activities. Another advantage is that these kinds of computer-based assessments can also be accomplished in real time, which allows for feedback to the learner also in real time, but again not in a manner that makes the learner aware that they are being assessed.

In the following section we describe examples of approaches to conducting this kind of embedded assessment: "stealth" assessment in commercially available games, web and project analysis in Scratch programming, and learning analytics methods to assess programming and learning in heavily documented environments.

A Digression on Evidence-Centered Design

Key projects reviewed here employ variants of a conceptual framework for creating learning assessment procedures known as Evidence-Centered Design (ECD). ECD makes formal and explicit both the process of designing assessment procedures and the procedures themselves. It provides language and concepts for systematically developing interlinked models of the design, analysis, deployment, and measurement aspects of learning assessment (Behrens et al., 2004; Mislevy et al., 2004; Mislevy & Haertel, 2007). The process of developing these assessments involves analyses of the target-learning domain in order to establish domain representations, categories, and features relevant to addressing assessment goals. This established what is called the *competency model* for a specific assessment and answers the question: What collection of knowledge and skills should be assessed? ECD also requires specifying the kinds of student behaviors and actions that will constitute evidence of learning, the *evidence model*, answering the question: What behaviors or performances should reveal those constructs? Finally, these two models are connected by a conceptualization of how the assessment will be implemented, the *task model*, answering: What tasks should elicit those behaviors that comprise the evidence?

While initial efforts to apply ECD have centered on the development of assessments for use in formal educational settings, more recent efforts have applied the framework to activities and settings that fall within the scope of this report. We will describe three examples: so-called stealth assessment of player learning in commercially available first-person, role playing video games; networked analyses for assessing learning in epistemic games; and learning assessment of an innovative, large-scale implementation of a game-design curriculum.

Stealth Assessment in Commercially Available Games

Levels: Individual
Valued Outcomes: Independent inquiry
Methods: Evidence-Centered Design

Stealth assessment refers to a methodology for integrating learning assessments into the structure of computer game-based educational activities in such a way that relevant data about the learning process can be gathered in real time without disrupting student engagement in the activity. The main assumptions underlying stealth assessment research are that: learning by active participation in the game improves outcomes; different learner attributes may be identified during game play; strengths and weaknesses of the learner may be addressed to improve learning; and formative feedback can be used to further support student learning (Gee, 2003; Shute, 2007, 2008; Shute, Hansen, & Almond, 2008; Squire, 2006).

In addition to providing a rich context in which to study learning, video games are a central focus of stealth assessment because, as activities instantiated through computer platforms, they afford convenient and reliable tools for continuously recording and analyzing student performance data in real time. Gathering this documentation is not generally a problem; however, making sense of the potentially massive amounts of data generated requires a principled conceptual framework. Furthermore, once this data has been interpreted there is the necessary and complex task of feeding this information back into the educational activity. For example, in computer game-based environments this information can be used to support real-time adaptive adjustments to the game in ways that promote and maintain player learning. In the long term, it is also important to communicate to both students and teachers what was learned from these assessments, so that the learning activity can be restructured for improved and sustained learning.

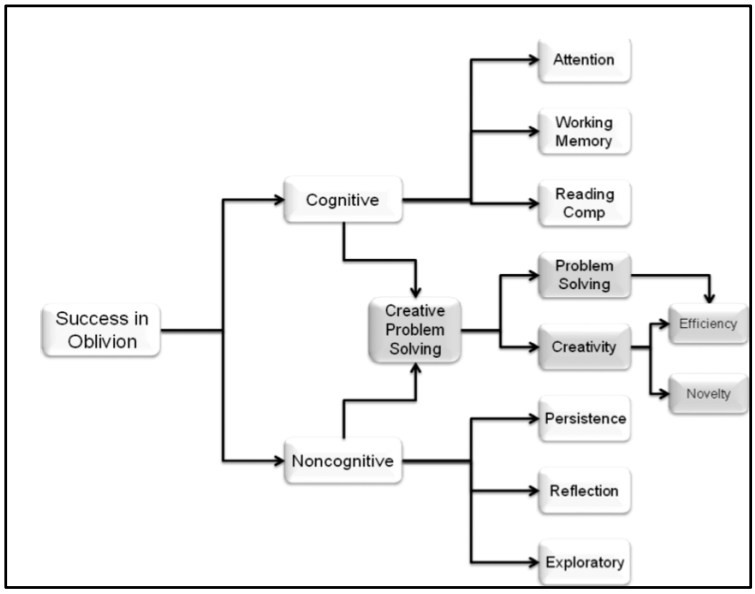
Over the past several years, Valerie Shute and her colleagues have pioneered stealth assessment methods (Shute 2011, Shute et al., 2009). In this section we draw on this corpus of work to explain the application of the ECD framework, described above, in assessing learning in video games. Games that incorporate stealth assessment need to elicit behaviors that provide evidence of the skills and knowledge being assessed, and they must afford principled interpretations of the evidence in terms of the aims of the assessment. As noted, ECD guides the researcher in developing a framework for collecting and analyzing data in terms of three interrelated models, the first of which is the competency model, which names the sets of knowledge and skills on which assessment inferences are to be based. We focus here on an example described by Shute et al. (2009), in which in which the researchers studied players' creative problem-solving abilities in *Oblivion*, a commercially available first-person, role-playing video game (*The Elder Scrolls IV: Oblivion*, 2006, by Bethesda Softworks).

Exhibit 8 shows a simplified competency model developed by Shute and her colleagues for modeling creative problem solving in the game. The model defines and situates creative problem solving in a restricted range of educationally relevant competencies that one could assess in *Oblivion* game play. Student-specific instantiations

of competency models are termed *student models*. The values reported in these models, like profiles or report cards, denote assessor beliefs about the student's level on each of the variables in the competency model.

Shute and her colleagues operationally define creative problem solving as a contextually specific "mental process of creating a solution to a problem . . . in which the solution is independently created rather than learned with assistance." By emphasizing problem solving according to this definition, assessment is focused on the novelty and efficiency of approaches to accomplishing tasks in the game.

Exhibit 8. Simplified Competency Model for the commercially available role playing game *Oblivion*.



The areas shaded in gray represent the variables used to assess creative problem solving. From Shute (2011).

The evidence model here defines the line of argument describing how and why player performances in a given task situation constitute evidence about competency model variables. Two questions guide the development of the evidence model: What behaviors or performances reveal targeted competencies? What is the functional (or statistical) connection between those behaviors and the competency model variable(s). Looking specifically at video games like *Oblivion*, connections between game observables and their corresponding competencies require that the evidence model include (a) scoring rules for extracting observables from players' game play indicators found in log files, (b) the observables (i.e., scored data), and (c) measurement rules for accumulating evidence from the observables, which are then used to update the student model variables.

In the present example (creative problem solving), novelty and efficiency link actions in the game with competencies related to creative problem solving. Efficiency is defined in terms of the quality and quantity of steps taken to solve the; novelty is defined in terms of the frequency with which certain actions are taken (i.e., low frequency actions are more novel).

The task model here serves as a framework for conceptualizing game-level situations that elicit player performances deemed to provide evidence of competency-related knowledge. Among other things, the model specifies what the player will be asked to do and what kinds of responses are allowed. In the specific case of stealth assessment, Shute and her colleagues refer to the task model as the *action model* in order to highlight the fact that the modeling is focused on defining player's action sequences and each action's indicators of success. Actions are defined as anything a player does to solve a particular problem contained within a scene in the game. Indicators are explicitly linked to each action, and these indicators can in turn be measured from the player's log file.

Assessing computer programming learning through the Scratch Community Website

Levels: Individual, group
Valued Outcomes: Collaboration, quality of product
Methods: Data mining (website)

Another area in which the documentation power of computers is being harnessed to assess student learning is for programming and digital media production. We return here to the example of the Scratch programming platform (see *Learning to Program in a Computer Clubhouse* discussion above). Maloney and his colleagues set out to describe which programming concepts students (8-18 years old) at a Computer Clubhouse used in their projects, how they used these concepts, and the extent to which the community of the Clubhouse as a whole increased its collective knowledge of computer programming over time (Maloney, Peppler, Kafai, Resnick, & Rusk, 2008).

Documentation consisted of both qualitative and quantitative measures. *Scratch* project summary files were exported and collected. These contained information about the number and kind of programming commands used together with records of the frequencies of stacks, sounds, and costumes used in the project. In addition, weekly fieldnotes of Scratch activities were gathered by university students, who attended the Computer Clubhouse as both researchers and Scratch mentors. Finally, interviews with Clubhouse members were conducted to document their impressions of Scratch and to obtain histories of their programming experience.

The summary files were analyzed for frequencies and distributions of Scratch commands in the corpus of projects studied. This allowed Maloney and his colleagues to develop a profile for Clubhouse members *as a group* of the programming commands in use over an 18-month period. They compared frequencies of programming commands included in projects from one year to the next, and analyzed the difference in percentages of projects containing targeted programming concepts. Significant gains were observed in use of five out of seven targeted programming concepts.

One finding that highlights the importance of the relationship between the social dynamics at the Clubhouse and what individuals in these settings do with these technologies is that the particular programming commands that were frequently used were commands that were important for creating programs that were seen as valuable by members. For example, in this particular Clubhouse games and animation were popular among the students. This in turn was reflected in the fact that many of the Scratch programs produced by the students included commands relevant for games and animations.

The university mentors selected to assist Clubhouse members with Scratch themselves knew little to nothing about programming. This was by design; with inexperienced programmers as mentors, Clubhouse members were expected to feel more empowered in their learning, even bolstering their programming expertise in situations when they would be asked to help mentors learn about Scratch (Kafai et al., 2007). In their study Maloney et al. (2008) argued that given the minimal role played by mentors at the Clubhouse in terms of any direct instruction in Scratch, the Scratch environment itself was likely most responsible for the learning that resulted.

In order to assess the power of the Scratch environment as a stand-alone programming learning environment, Dahotre, Zhang, and Scaffidi (2010) studied Scratch programs created by youth who did not attend Computer Clubhouses. To do this they used a “screen scraper” program to download from the Scratch community website 100 random animation projects, their corresponding computer code, and user comments. The HTML for the repository web page was also downloaded in order to assess usage statistics (counts of views, comments, downloads, and visitor re-mixes) for each animation project. These projects were then assessed in terms of three learning goals envisioned by developers as being facilitated by Scratch: learning of technical programming skills, social skills for collaboration, and socio-technical re-mixing skills related to adapting existing programs and community resources to produce new programs.

The projects that Dahorte et al. (2010) evaluated from the community website included key programming primitives at the same levels as the projects that were developed under partial supervision of undergraduate mentors (e.g., Maloney et al., 2008). In order to assess these projects against the general population, the authors drew on findings from methodologically similar studies based on other repositories of end-user programming code (e.g. spreadsheet programming and macros). They found that the Scratch programs showed comparable or higher rates of use of programming primitives. To assess collaborative and re-mixing skills, the authors studied the comments posted by users to projects on the website and project download statistics, respectively. While the majority of comments posted were helpful critiques, they did not indicate that students were actively collaborating with one another on projects. Regarding re-mixes, the authors found that the relatively low levels of remix activity were comparable to those found in the general population.

More recently, Scaffidi and Chambers (2011) extended the work of the above two studies to examine the extent to which the Scratch environment, via the web community, facilitated the development of social skills or elementary programming skills. To this end they studied data from over 1000 Scratch projects posted to the community website by 250 randomly-selected users. Drawing on existing literature, Scaffidi and Chambers adapted four models of end users to create a framework for investigating skill progression in Scratch. Two models addressed

questions of programming skill related to the use of programming primitives, one examined social skills, and the last concerned programming efficiency based on the number of lines codes produced and the amount of time taken to produce them.

Learning Analytics in computer programming

Levels: Individual
Valued Outcomes: Independent inquiry
Methods: Learning Analytics, data mining (programming code)

Paolo Blikstein and his colleagues at Stanford's Transformative Learning Technologies Lab have developed a number of automated learning analytics approaches for make quantitative assessments of the kinds of open-ended, non-scripted learning activities that, as this survey of the literature demonstrates, have been previously largely qualitative. One example of this work concerns the assessment of students' programming know-how.

Blikstein (2011) examined assignments completed by nine sophomores in an undergraduate programming class. The students were asked to choose a specific scientific phenomenon and write a program that modeled the phenomenon. Student programming logs were collected using the *NetLogo* (Wilensky, 1999) programming environment. *NetLogo* is capable of logging to an XML file such user actions as key presses, button clicks, and changes in programming code.

Blikstein developed a special configuration file for specifying and logging the targeted programming actions. This file was distributed to the students who were given instructions for enabling it so that log-files could be collected and then processed, coded, and analyzed. The analyses focused on identifying the coding strategies that the students employed. Once identified these strategies were combined with survey information about the students' previous programming knowledge to determine coding profiles for each student (e.g. copy and pasters, self-sufficients, or mixed mode). The broader aim here was to draw on the students observed programming patterns to improve the design of teaching and support materials and strategies, as well as to identify critical points in the process of writing software applications where support interventions might best be included.

Ethnographic Studies of Online and Gaming Communities

The following two studies represent a useful approach to research and assessment of online communities in which the researcher acts as a member of the community or otherwise studies the online community as an ethnographer might study a traditional face-to-face community. Insofar as informal learning increasingly occurs today in such online communities, methods of documentation and assessment developed within them are important for the field.

Lineage, World of Warcraft, and Second Life

Levels: Individual, group
Valued Outcomes: Collaboration, independent inquiry, social resources
Methods: Cognitive ethnography

Steinkuehler (2008) has described a program of research for studying learning in massively multiplayer online role-playing games (MMORPGs) such as *World of Warcraft* (WoW), as well as non-game online virtual worlds like *Second Life*. The first phase of the research involves conducting a cognitive ethnography (Hutchins, 1995) of the game and its community of players with the aim of identifying the forms of cognition and learning that lead to successful gameplay. In her own research, Steinkuehler studied the fantasy-based MMORPG, *Lineage*. She engaged in 28 months of participant observation. This involved the study of naturally occurring game play documented through video and fieldnotes, interviews with players, and the archiving and analysis of game-related player communications (e.g., emails, chatroom and instant message conversations, and discussion board posts) and community documents (e.g., fan websites, community-authored game fictions). From this work Steinkuelher identified five broad categories of social and intellectual practices that characterize successful MMORPG play: (1) collaborative problem-solving, (2) digital media literacy practices, (3) informal scientific reasoning, (4) computational literacy (e.g., modding), and (5) cultural mechanisms for learning, including reciprocal apprenticeships (Steinkuehler, 2004) and collective intelligence (Steinkuehler, 2006). See Steinkuehler (2005) for a full survey of the practices identified.

Note that for players of MMORPGs there are normally two intersecting online communities: the in-game community of interactions with other players during actual gameplay, and the out-of-game online guilds and associations in which information about the game and game planning and strategies are developed and exchanged.

In the second phase of research, Steinkuehler drew on the social and intellectual practices identified in the first phase and applied them as guiding themes for conducting targeted empirical investigations of *World of Warcraft* and *Second Life*. These studies yielded observations about collaborative problem solving and the multiple literacies that players develop as a result of participation in the game and the game community. Steinkuehler notes important parallels between the kind of collaborative problem solving required in games like *World of Warcraft* (e.g., group-organized “raiding”) and those required in contemporary workplace environments. She also reports on the high degree of motivation that players show for developing multiple literacies, including computer

literacies (e.g., creating mods) and practices that overlap with more traditional literacies (e.g. writing prose in the form of fan fiction). For an ethnographic study focused on collaborative problem solving in *World of Warcraft* see Chen (2009).

Whyville

Levels: Individual, group
Valued Outcomes: Collaboration, independent inquiry, social resources
Methods: Connective ethnography

Online virtual worlds have become pervasive as educational tools both in and out of schools. Regardless of the setting, implementing these worlds and assessing their potential learning benefits remains a challenge. A key question raised by environments in which learners inhabit and travel between virtual and physical spaces is how this learning can be traced. To examine this question, Fields and Kafai (2007) conducted a connective ethnography of a summer after-school program in which 4th-6th graders used an online virtual world, Whyville.

Citing work by Hines (2000) and Leander and McKim (2003), Fields and Kafai define connective ethnography as an approach that focuses on examining how everyday practices work to segregate and blend together social spaces. These processes are studied by tracing the flows of objects, texts, and bodies and how these flows contribute to the construction of boundaries within and between virtual and physical spaces. Within this framework, Fields and Kafai adopted a strategy of identifying forms of insider knowledge about the game that would afford insights into the kinds of learning that players experience in and across the virtual world and the physical space of the after-school center. This knowledge needed to be traceable, discoverable by players through trial and error, and observable in and out of the virtual world (Fields and Kafai, 2007, p. 200). Understanding the particular insider knowledge that they studied—teleporting—requires a brief introduction to the basics of the game.

Whyville is a multi-user virtual environment for 8 - 16 year old youth in which users are positioned to engage in a variety of science-themed activities. Each time a player successfully completes an activity they earn virtual currency, which they can then use to buy and design features of their in-game avatar, and other goods useful within the world of the game. Social interaction in the game is mediated through “ymailing” (email within the *Whyville* universe) and avatar-to-avatar chat. Avatars interact in over 30 different settings, each of which can be accessed via a pull down menu (see Exhibit 9).

Exhibit 9. From left to right: the pull-down destination menu, and two of the 30 destinations where players can visit and interact, the beach and the moon. From Fields and Kafai (2007).

Not all of the locations that players can visit are visible in the menu (e.g. the Moon, Jupiter, and Saturn are not). These "secret" sites are only accessible by "teleporting," an action that can only be accomplished by typing "teleport [place]" in the chat bubble above the avatar's head. Teleporting is not observable in the game because avatars are transported to the new location before the typed commands are displayed, preventing other players from knowing where and how the avatar disappeared. Where these secret locations are and the ways of getting there are bits of information that can only be obtained through conversation with other players or from *Whyville* "cheat" websites. Consequently, these secret locations gain a special status among *Whyville* insiders as social hang-outs free of newbies (new, relatively naïve players).

To document learning over the course of the 9-week program, Fields and Kafai adopted a data multi-streaming approach. Daily fieldnotes were taken to document the overall activity at the after-school center while video was used capture interactions of small student groups as they played while clustered around a computer. Online tracking data was recorded for players' chat and the locations of their avatar throughout the game. Finally, individual post-activity interviews of the participants were conducted.

All text documentation was combed for instances in which teleportation was mentioned. This, for example, allowed for a mapping of each player's history of learning and using the teleportation function. For example, chat transcripts for each student were studied to identify the first time that students discussed teleportation in the game and when they first used the teleportation function. This information was in turn triangulated with fieldnotes and video data in order to paint a more detailed picture of the contexts in which players learned and implemented the teleportation function.

Fields and Kafai found that for the student players there was a seamless integration between online and offline interactions. What happened in *Whyville* was just as important as what happened in the after-school center. For this reason, Fields and Kafai propose referring to activity in these worlds as "synthetic play," as in the synthesis of online and offline worlds.

Additional Studies of Interest

We include here more detailed accounts of two additional research projects that have developed relatively sophisticated approaches to documentation and assessment of settings in which digital games provide key learning opportunities. In each case these projects draw on and elaborate the Evidence-Centered Design framework discussed earlier.

Epistemic Games and Epistemography

Levels: Individual, group
Valued Outcomes: Independent inquiry, know-how, collaboration, social resources
Methods: Evidence-Centered Design, Epistemography, Epistemic Network Analysis

The research we describe here is drawn from the work of the multi-institutional Epistemic Games Research Group led by David Williamson Shaffer and which includes researchers, faculty, staff, and students of the University of Wisconsin at Madison, the University of Maryland, the University of Memphis, the Massachusetts Audubon Society, the Danish School of Education, and the Open Universiteit Nederland. Shaffer and his colleagues study digital learning systems: systems constituted by a theory of learning and corresponding methods of assessment, which are linked into an evidence-based, digital intervention (Shaffer et al., 2009). The digital learning systems that Shaffer and his colleagues develop and study are known as “epistemic games”.

Epistemic games are activities through which students learn the knowledge, skills, practices, and values of a particular professional domain. These activities are mediated through a variety of digital technologies including virtual world simulations of the tasks typically engaged in by professionals in the target field, as well as computer applications used by professionals themselves in their everyday work. Players take on the role of apprentices-in-training in such professions as city planning, science journalism, and engineering. Learners are guided in these worlds through the deliberate inclusion of game tasks that ask learners to reflect on their work. This can involve interactions with peers and/or with actual professionals in the relevant fields. These interactions are designed by the game creators and researchers to model the kind of mentoring that professionals experience in their training and socialization.

One example of an epistemic game is *Digital Zoo*. In *Digital Zoo* players adopt the role of biomechanical engineers given the task of using specially designed software to create virtual objects and creatures for an animated film. Design specifications about the desired appearance and movement of these objects and creatures are delivered to the players who are then tasked with creating these objects. Both through design elements of the activity as a whole and through the participation of knowledgeable mentors, the players are guided in following the roles and rules of engineering design based on research examining how real engineers-in-training learn to design.

The epistemic frames hypothesis (Shaffer, 2006) is a theory of learning on which the design, implementation, and study of epistemic games are based. The hypothesis proposes that any community of practice (e.g., biomechanical engineers) has a culture with identifiable structuring features including: the things that people

within the community do (skills); the understandings that people in the community share (knowledge); the way that members of the community see themselves (identity); the beliefs that members of the community hold (values); and the warrants that justify actions or claims as legitimate within the community (epistemology). For more details see Shaffer, 2007; Shaffer et al., 2009.

The skills, knowledge, identities, values, and epistemologies of the community of practice collectively constitute that community's epistemic frame. The epistemic frame hypothesis argues that participants engage in specific forms of training and socialization to become members of these communities. These processes lead participants to internalize these epistemic frames which they can in turn deploy as tools for engaging with the world from the point of view of a community member (Shaffer, 2006).

Evidence-Centered Design in the Epistemic Games Project

A critical step in the implementation of an Evidence-Centered Design approach is to conduct domain analyses in order to identify the target learning to be assessed (the competency model and the individual instantiation of this model, the student model). From an epistemic games perspective, the development of the competency model is guided by the epistemic frame hypothesis where the model is understood to be constituted by those elements of the epistemic frame that one seeks to have players internalize.

How these elements are organized and engaged within the course of the activity is based on what is known about the professional practica in place for training and socializing members of the professional group. Drawing on Schon's research in professional development (1983, 1985), Shaffer and his colleagues argue that epistemic frames develop and are passed on through these practica: "Professional practica are environments in which a learner takes professional action in a supervised setting and then reflects on the results with peers and mentors. Skills, knowledge, identity, values, and epistemology become more and more closely tied together as the student learns to see the world using the epistemic frame of the community, as happens in capstone courses in engineering, internship and residency for doctors, or almost any graduate program in the sciences" (Shaffer et al., 2009).

Defining the elements of the epistemic frame to be replicated in the epistemic game requires that the researchers perform an *epistemography*. An epistemography is an in-depth study of the participant structures in a specific practicum environment for the purpose of identifying learning processes that facilitate a novice professional's development of a particular epistemic frame (Shaffer, 2005). Participant structures are the forms of action and interaction, including reflective practices, typical of profession.

In the specific case of the *Digital Zoo,* the designers conducted an epistemography of an undergraduate engineering design course (Svarovsky & Shaffer, 2006a, 2006b). Using a variety of ethnographic methods (field observations, interviews, focus groups, artifact archiving and analyses) the designers followed a student design team as they worked to develop a biomedical device for an actual client. Qualitative data from student design notebooks as well as from regular design meetings were collected. These data were then analyzed using an initial coding scheme developed from descriptions of practice in the literature (Burghardt, 1999; Dym & Little, 2000) and

the definition of an epistemic frame (Shaffer, 2006a). Adopting a grounded theory approach (Glaser & Strauss, 1967; Strauss & Corbin, 1998) the codes were iteratively refined.

Analyses revealed that design meetings, design notebooks, and meetings with clients were effective means for reflective activity. In particular, the design meetings and design notebooks proved to be rich sources of epistemic statements about engineering that were highly correlated with references to engineering skills, knowledge, and values. For the researchers this suggested that these activities played an important role in the initial development of students' engineering epistemic frames, and therefore would be incorporated as important activities in the epistemic game.

A storyboarding procedure (a *frameboard*) is used to design and implement the systematic presentation of these activities (i.e. participant structures) in the game so that students engage in them in a reflexive manner. From the perspective of ECD, the frameboard functions as the task model. Each activity in the frameboard describes the relevant participant structures, defining them in terms of the features of the epistemic frame that are required to engage this structure at the particular point in the game, and with respect to the forms of evidence necessary for claiming that these features have been developed and properly deployed.

Epistemic Network Analysis

Shaffer et al. (2009) argue that the kind of professional perspective that students develop from participation in epistemic games should not be thought of solely as a collection of the relevant skills, knowledge, values, identities, and epistemologies of the profession. It is also important to think of these elements as constituting a network of conceptual, practical, moral, personal, and epistemological relationships. It follows that assessing student learning requires not only carefully documenting instances of students engaging the different elements of the frame, but also characterizing the strength of the relationships among these elements for each student.

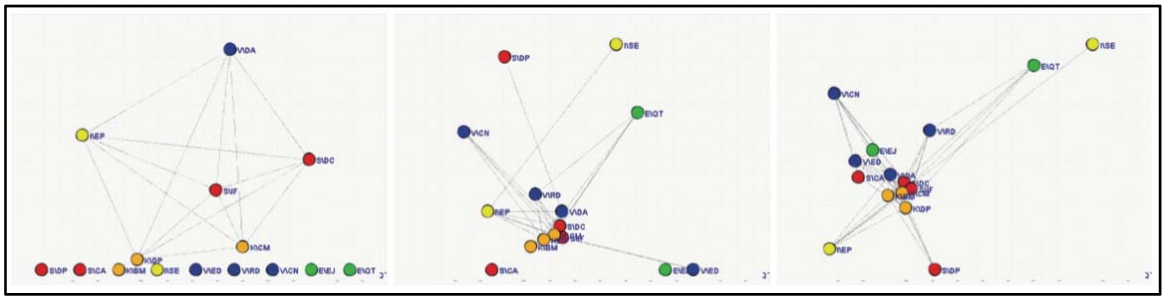
One method that the Epistemic Games Research Group uses to study these elements and their relationships is Epistemic Network Analysis. The approach involves a combination of social network analysis (Brandes and Erlebach, 2005), concepts from frame analysis (Goffman, 1974), and data visualization techniques. To explain the application of social network analysis in assessing learning in epistemic games, Shaffer et al. (2009) draw an analogy between using the technique to analyze interactions at a party and relations among the elements of an epistemic frame. One way to examine the relationships among people meeting for the first time at the party would be to document at intervals who is interacting with whom. Assuming that a closer relationship develops among those who spend more time in the same conversational group, the social network that develops can be quantified by summing, for each pair of persons over the course of the party, the number of times they are found in the same conversational group. Analogously, we can think of the people interacting at the party as the elements from an epistemic frame for one player in an epistemic game. The task then is to trace over time those elements of the epistemic frame that the player uses.

To understand how these networks of frame elements are visualized we return to the *Digital Zoo* epistemography. The documentation collected included player design notebooks, documents, reports, and other work products

produced during game play. Player-mentor interactions were also recorded. For each player these records were then organized into play histories. Using working definitions of the five engineering epistemic frame elements (derived from the earlier qualitative coding process), these histories were studied to identify instances of these elements.

Shaffer and his colleagues use network graphs as tools to create visualizations for individual players of the epistemic frame elements that emerge at specific moments in the individual's play history. These graphs not only show which frame elements the player develops over time, but also the nature of the relations among those elements. In a cumulative graph the length of the line connecting frame elements indicates the frequency with which each pair of elements co-occurs in a strip of activity: the closer the nodes, the more often they were linked during the game. Using this same approach, graphs can be created for players at any point in the game and these in turn can be used to perform developmental analyses of epistemic frames (see Exhibit 10).

Exhibit 10. Network graphs from different slices of the play history for one player in *Digital Zoo*



From Shaffer et al., 2009.

Statistical techniques can be applied to run a variety of analyses on these epistemic network maps, providing insights into the development of players' epistemic frames. For example, linkages among nodes in a network can be measured at different time intervals to assess changes in the overall strength of association of the network (weighted density). These measures can in turn be applied to determine changes in an epistemic frame over time, and to associate those changes with specific elements of game play.

Evaluating Quest to Learn

Levels: Individual, Group, Project
Valued Outcomes: Independent inquiry, collaboration,
Methods: Evidence-Centered Design, Evidence-Centered Evaluation, content analysis

The following example highlights assessment at all three levels regarding independent inquiry and collaboration. While this example involves assessment and evaluation of what would be typically considered a formal learning

environment—a public school— both the models of assessment and evaluation applied (Evidence-Centered Design and Evidence-Centered Evaluation) and the unusual nature of the curriculum around which activities at the school are organized (video-game design) make this project also relevant for designing and implementing assessments of informal learning activities.

Quest to Learn (QL) is an innovative New York City public school whose curriculum is organized around the principles of game design. It is the first school of its kind, a product of the growing movement in education and education research to study and apply principles of game design to create motivating learning activities; see for example Shaffer and Gee (2012). Given the school's unique curriculum, one obvious question that emerges is how to assess learning in an environment that has never been formally assessed before. Shute and Torres (2012) took on this challenge by adopting a two-pronged approach: to assess learning at the individual level and to evaluate the success of QL project overall. They used Evidence-Centered Evaluation, described below, to evaluate and model the goals of QL, and Evidence-Centered Design (previously discussed) to develop and implement assessments of individual student performance.

Shute and Torres describe Evidence-Centered Evaluation (ECE) as an extension of Evidence-Centered Design (ECD). Both approaches are designed to guide the researcher and practitioner in specifying the structures and supporting rationales for the evidentiary argument of an assessment. Argument structures encompass, among other things, the claims (inferences) one wishes to make, the observables (performance data) that provide support for those claims, the task performance situations that elicit the observables, and rationales for linking them all together.

Like ECD, the procedure for applying ECE involves first defining a set of three interrelated models for objectives, evidence, and data. The objectives model identifies the goals of the organization. It is a description of what is of value to the school and its extended community. As such it guides the development of metrics and criteria for evaluation of what works, or not, and why.

The first step in modeling the QL's goals is data collection. Over a period of seven months Shute and Torres conducted semi-structured interviews, surveys, focus groups, and observations with and of the relevant stakeholders including teachers, students, administrators, and curriculum designers. They also studied the school's design documents (see Salen, Torres, Wolozin, Rufo-Tepper & Shapiro, 2010. Dominant themes were identified in the interview data using content analysis. These themes where then summarized in a document that was shared with all of the participants. The participants were then given a survey to record their responses to and feedback about the document. After review of these surveys, focus groups were organized as a means of further refining the themes. The QL objectives model was developed based on this set of themes (see Exhibit 11).

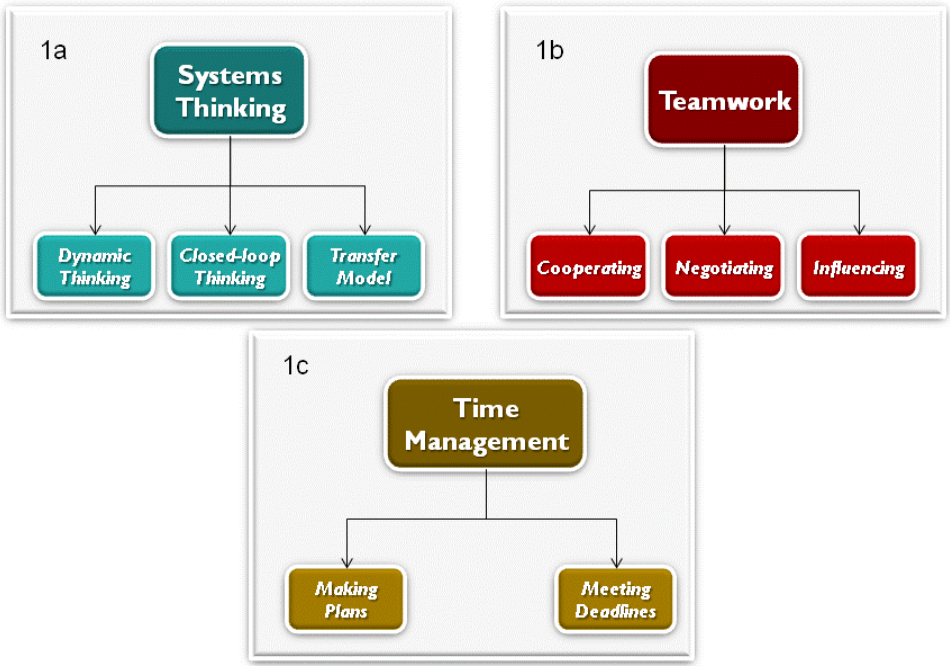
Exhibit 11. Quest to Learn Objectives Model

Assessment of student-level variables centered on three competencies valued by the QL community: systems thinking, teamwork, and time management. In line with an ECD approach, competency models for each variable were developed. Shute and Torres derived these specific models from the research literature, and these models were in turn used to develop corresponding assessment instruments (Exhibit 12). These instruments, which were administered to QL students at six-month intervals, were just one component in a battery of assessments that mixed qualitative and quantitative measures. A qualitative observation protocol was also implemented in which two independent observers conducted bi-weekly observations in classrooms, the cafeteria, and the school's after-school program. The focus of these observations included: (a) thinking skills afforded by each node of the model (e.g., data on systems thinking, observations focused on evidence related to dynamic thinking, closed-loop thinking, and skills for transferring learned models across multiple situations and settings); (b) type and frequency of specialized language used within each domain/node; (c) kinds of social activity evident per node; and (d) other relevant information including learning tools used, identities afforded by each site, artifacts produced, shared norms, physical (or virtual) space, and time allocated per activity in a given location (Shute and Torres, 2012).

Mindful of concerns about the relationship between participation in the innovative activities at QL and achievement in traditional academic domains, Shute and Torres also planned to compare performance on state mandated standardized tests between QL students with groups of demographically matched students at two New York City public schools.[2]

[2] Our description of Shute and Torres' methods in this section is based on published work of an in-process, 18 month study in which they report only preliminary findings related to measures of systems thinking, team work, and time management (see Shute & Torres, 2012).

Exhibit 12. Three competency models for student assessment at Quest to Learn



From Shute & Torres (2012: 99)

Summaries of the Expert Meetings

From June 2011 through February 2012 three expert meetings were convened, with a total of 25 senior researchers participating (see list of names in Appendix A). All brought rich experience with project design and development, internal project documentation and assessment, and in many cases also participation in external assessments of research across a wide variety of informal learning settings. We included scholars with experience in conceptualizing and theorizing learning and assessment as well as those whose primary focus is conducting empirical research.

At the first meeting, held in San Francisco and hosted by the Exploratorium (a leading informal science learning institution), the core agenda was defining the scope of this project in terms of the key features which define informal learning settings of interest and the kinds of settings to be included or excluded. We also discussed how to define an appropriate unit of assessment for learning in non-formal settings. The group viewed videos supplied by participants of examples of learning in several different settings and projects and sought to connect our conceptual discussions to these concrete examples. This procedure proved very helpful and we repeated it at all the meetings.

We also sought and received from meeting participants (at all three meetings) suggestions regarding significant research projects, both completed and in-progress, names of other leading researchers to consult, and citations to reports and published literature of relevance. This resulted in a bibliography included in Appendix B, which is far more extensive than the list of studies chosen for review in the previous section and also in a list of other relevant resources, including websites (see Appendix C).

The second meeting was hosted at Northwestern University near Chicago and focused on both major conceptual issues and on research on learning with digital media. Issues of the nature of learning communities, the affective dimensions of learning, learning by groups and projects, new modes of documentation, and trajectories of learning over time were prominent in the discussions. There was considerable discussion of assessment methods and techniques and on the topic of “badges” as a means of recognizing achievement in informal learning.

The third and final meeting was held at the University of California, San Diego. The focus of this meeting was a review of the basic conceptual structure and plans for the report, some discussion of possible conclusions and recommendations, and more discussion of learning communities, the goals of assessment, badges and other recognition systems, and appropriate diversity of kinds of valued outcomes.

At both the second and third meetings, participants were provided with summaries of the prior meetings and discussions as a result were frequently cumulative. The third meeting, in particular, reflected on and refined proposals from the earlier meetings. Overall, there was a broad sense of agreement on key issues, with a few limited exceptions. Later meetings tended to add to and clarify proposals from earlier meetings, and across all

three meetings there was a strong sense of the importance of building a sophisticated and critical consensus in the field.

We provide below our internal summaries of the highlights of the meeting discussions, edited to limit redundancy with what has been incorporated elsewhere in this report. On the project wiki we also include more detailed minutes of the meetings and details of the discussion: http://documentinglearningworkshop.wikispaces.com/

San Francisco Meeting Highlights

A key issue discussed was that of **the unit of assessment.** The consensus of the group was that this unit is:

> *A system over time, which includes individual learners, but also other participants, mediating tools and semiotic media, and local conditions directly relevant to and supportive of the learning activities.*

It needs to be extended to wider contexts that make the setting of learning possible institutionally, but with decreasing detail as relevance to the specifics of learning trajectories decreases.

This was later applied to the question of meta-data or backstories for **video records of learning activities**. A key question, answers to which may vary by type of setting or type of activity, is:

> *How much and what kinds of information are relevant to identifying valued learning and the specific aspects of activity that support it in the video?*

A suggestion for estimating the role of background information is to have a group view the video first without background, and then again with partial background, and a third time with much more complete background.

It was suggested that in any records of learning activity over time, some valued learning may be more readily visible, to more observers and with less detailed analysis or less experienced professional vision, while other instances may be less readily visible.

It was also agreed that different kinds of learning may become visible when records of learning activities are studied over longer vs. shorter time periods.

Another key proposal was to identify and survey research and assessment in and across a variety of **types of settings**, including:

- *Learning in the home* through everyday activities or activities not specified by the requirements of some other educational institution (example: doing and discussing mathematics during home remodeling)
- ***After-school programs***, where activities are not directly meant to serve school-based academic functions (example: playing an educational computer game and making innovative use of it for fun, with ancillary learning)
- ***Community center programs***, where activities are negotiated between learners and providers, and which may have specific learning objectives but changing approaches to the goal (example: tele-mentoring and use of computer simulation of electric circuits, together with an onsite coach familiar with the student, but not responsible for the content)
- ***Museum-based programs***, where visitors can choose to manipulate hands-on materials in the context of questions and explanations of phenomena observed/produced (example: young visitors connecting a battery to various electric devices to see the results of completing a circuit, with a coach, and showing the results to a parent; a group of young visitors extracting insects from a bag to feed to a pet as part of a longer term project, and one overcoming a reluctance to touch the insects)
- ***Online communities and forums***, where participants ask and answer questions on a specific area of competence or expertise and evaluate one another's answers or contributions, and where they may also engage in joint activity in a virtual space or mediated by tools and social interactions in that space (example: learning to build in Second Life; "theory-crafting" to identify technical characteristics of computer games by systematically playing many options within them; modding in World of Warcraft; raiding as joint play for a goal).
- *Team sports*, both live action and "fantasy" or virtual-world mediated
- *Crafting communities*, online and offline, such as Ravelry (knitting, etc.)
- *Cooking communities*, online and offline.

NOTE: It was ultimately decided, in order to focus the scope of the report in relation to available project resources, to include only the four types of settings identified by bold text. by asterisks in this list.

Issues of Value and Significance

By what criteria do we decide that some learning is valuable?

It was recommended that the first criterion be that there is evidence of value for the participants, e.g. through the length of time they focus on a task or activity, their reluctance to leave or end the activity, displays of intense or positive affect during the activity, comments on the activity during and after, explicit elicited evaluations.

Beyond this there are additional criteria which may be applied, primarily: the judgment of expert educators or others, such as parents, regarding what is of value to the learners and/or to society; evidence of consequentiality of learning for other, conventional academic activities (e.g. increased interest, increased participation, more positive affect, more effective completion of tasks, ability to teach content and skills to others, ability to solve problems collaboratively, etc.)

Issues of Documentation

Video alone is often not sufficient documentation of learning activity because of the inferences that need to be made to identify valued learning. Ideally *video should be supplemented by fieldnotes from participant observation or observant participation, interviews with participants, relevant histories of the setting and of the participants*.

> *The unit of analysis for documentation should be activity in a setting over times long enough to show: origins of participation, evolution of the activity, learning in the activity, consequentiality of learning for some other activity.*

An individual episode captured on video may be significant as part of such a longer trajectory of learning and development, and/or as an instance of a frequently repeated pattern of learning and activity across other cases in the same or similar settings.

Issues of Definition and Scope

“Informal learning” is at best a shorthand for a more complex combination of specific features which need to be identified. These features may in principle occur in both school and classroom-based learning and in other settings, but in different combination and to different degrees. Each setting and perhaps each kind of learning activity will tend to have a particular combination and degree of each feature.

Some of the features noted as relevant include:

- Voluntary participation
- Enjoyment of the learning activity for its own sake
- Intense engagement with tasks
- Flexibility in goals and re-purposing resources
- Unpredictability of some significant learning outcomes
- Improvisation and innovation within and concerning the activity
- Commitment over time
- Continuing voluntary participation despite setbacks
- Relatively equitable power relations in negotiating goals and means

Various literatures may name activities or settings where these features are present, dominant, constitutive, or highly significant ("interest-based learning," "free-choice learning," "nonformal learning," " learning in passion communities," etc) as well as making distinctions among these based on role relationships or types of institutional goals and constraints.

Highlights of the Northwestern Meeting

In order to do **assessment across settings**, cumulative or comparative, we need to identify what that matters remains the same across the settings. Many common assumptions about this are flawed. E.g. identities change across settings, modes of learning do, purposes do, so do what is valued from the learning activities.

Communities are not bounded, fixed entities, but abstracted from flows of practices among participants in many communities. Learning cannot be defined as progress toward mastery in a community given this fact.

Assessment needs to be based on **longitudinal, ethnographic records,** e.g. collections of material objects and semiotic products with in-progress versions, over time.

Affective engagement may be one index that is common across settings and so important for wider assessment. Includes commitment, persistence.

The **STEAM approach** offers additional grounds for assessment, e.g. in the design of artifacts and re-mix products. [Note: STEAM refers to the addition of Arts and humanities perspectives to learning in STEM – science, technology, engineering, and mathematics.]

It is not just individuals that learn, but systems. We need to document the ways in which a **system of learning tools and learners** improves its effectiveness by various criteria. *Individual learning may be functionally definable only in relation to the wider learning of the ecosystem*.

Changes in **social networks** and the distribution of practices across networks can be evidence of system-level learning (including individual learning).

Badges are useful as a community-internal tool to recognize accomplishment, but are dangerous potential tools of social control if they are publicly collected and endorsed by large-scale authorities.

Badges may represent one instance of a valuable concept: **crowd-sourcing of assessment,** within a community of expertise. In order to assess individuals in depth, or small groups, project groups, etc., you need the power of large numbers of assessors who share some basic values and expertise.

The **evidence of learning** at level N in a game-like model is the player's performance at level N+1. [Note: "level" here does not mean level in the sense used elsewhere in this report (i.e. level of organization, unit of analysis) but rather level in the sense of "difficult level" or level of advancement in a game, as in the term "leveling up" to a higher level of challenge and resources.]

Methods are needed to insure that when an important learning event is recognized retrospectively, we can capture it and what immediately preceded it (cf. "TIVO" camera method). This suggests the value of comprehensive or always-on **video recording.**

Two key questions for further research and discussion were raised:

> *What methods of assessment are most useful for the purpose of improving the learning environment (system, tools) itself?*
>
> *What quantitative measures in a local setting are most useful as data for assessment across settings and across projects?*

Highlights of the San Diego Meeting

Aims of the Report: To provide leadership within the community in helping define good practice in assessment of learning in informal settings, to identify which methods may be most useful for assessing which kinds of learning outcomes, to provide examples and exemplars, and to make recommendations for future project assessments and needed research on assessment.

To this end, it was recommended that we *organize the literature review* and discussion by types of valued outcomes, and within outcomes by useful methods of documentation and assessment, with examples and citations.

[Note: in practice this system of organization proved impractical, so the literature review is organized instead by settings and types of projects, but we will include on the project Wiki a table identifying studies by outcome types and assessment methods]

New and promising approaches: Data mining, machine learning analysis, richly instrumented spaces, tracking learning across programs and settings, collecting and sharing data about learners across programs, longitudinal assessment over periods of 5-10 years, digital storytelling as a mode of documentation, spatial tracking, and agent-based modeling were all identified in the course of the discussions.

Critiquing and improving basic concepts: the key concepts of learning, knowledge, assessment, evaluation, documentation, engagement, enjoyment, agency, and outcome were all discussed in a review of the discussions in prior meetings in order to make them more useful for discussions of informal learning activities. The consensus formulations are given in this report in the Introduction.

Understanding Learning Communities – in order to adequately document, assess, evaluate, and research them

- Learning communities differ in their basic goals and values, strategies, organization for learning, and in the roles they make available for members and the new niches members may create for themselves in the community

- Learning outcomes for individuals, groups, and communities must be documented, assessed, and evaluated relative to differing goals and values (e.g. those of individual participants, those the community considers appropriate to various roles, those of the group and community itself, and those of external communities, e.g. sponsoring organizations)
- Learning goals for many communities focus on the drivers of learning more than on the learning of specific content; they aim to improve motivation, engagement, and enjoyment; to broaden areas of interest and expand zones of comfort; to improve skills of self-guided learning, sustained learning, and collaborative learning
- Some communities and some roles may emphasize activities which participants already value and enjoy, while others may draw participants out of their comfort zones to expand their know-how and its range of mobilization
- Informal learning communities often differ from school-based learning groups in terms of: age-mixing, flexible pacing, division of labor and goals according to role, unpredictable learning outcomes, self-guided learning, voluntary participation, etc.
- Some learning communities are organized in such a way that groups and the community as a whole learn and change over time, while others are mainly organized to facilitate only individual learning

Recognition Systems / Badges

Badges were offered as a means of recognizing achievements (a) made in non-traditional settings, and (b) made in relation to kinds of know-how for which there may not be formal recognition systems. Badges used within communities, awarded by the community and its members, serve both to recognize achievement and to make those who have made achievements recognizable to others as potential sources of expertise and assistance. Badges can also help to define ladders of participation, which indicate for the learners the existence and nature of higher levels of skill.

The value of a badge depends on the reputation of the communities and procedures by which it is awarded. A recognition system must be integrated with a reputation system if it is to operate outside specialized communities.

Badges are an example of the **crowd-sourcing of evaluation**. Within a community, this mode of evaluation and recognition may implement community or group consensus, rather than an exercise of power, and so avoid some of the resistance-based invalidation of other forms of evaluation. But if a wide-scale system of endorsement of badges or badge awarders (e.g. by government authority) prefers its own criteria of value to those of the awarding communities, it could undermine the authenticity of badges as endogenous evaluations and trigger the same reactions and invalidation seen with traditional external power-based evaluations (e.g. grades, standardized testing). This could include badge-seeking without engagement, badge “inflation” lowering the perceived value and use value of badges, lowering of criteria for awarding a badge, efforts to obtain badges without durable, mobilizable learning, etc.

Reputation systems are also at risk of manipulation or degradation, as we see today with the relatively low standards for accreditation in the for-profit sector of higher education, the purchasing of institutions with already accredited programs, the likelihood of bribery, etc.

It is not clear what the likely effects on an open badges system would be if badges were awarded by for-profit institutions or organizations as a source of revenue. This should be a major concern.

Localized vs. Inter-operable Assessment and Documentation

Some methods of documentation and procedures for assessment produce value insofar as they are designed specifically in relation to the goals and practices in a particular community. Other methods and procedures can be used equally well across different projects, and so support comparisons and generalizations. It is important to balance these approaches in relation to the goals of improvement and generalizable knowledge.

For example, a coding scheme to identify a particular category of events in a video archive documenting a project may either focus on types of events that are highly specific to that project and its goals and which may not be in evidence or be relevant to any other project; or it may try to focus on types of events that are very likely to occur across other projects and be relevant to a wide range of goals. It seems desirable to try to include both kinds of focus in assessment.

This example points up the balance between the goals of assessment (improvement) and the goals of research (knowledge relevant to future design). But the relation is not as simple as might be imagined. In-depth assessment of a project on its own terms may be of enormous value for research purposes insofar as it produces knowledge about how particular outcomes were achieved, what worked and what didn't, what was sustained over time and what got changed, etc. This knowledge *can potentially be combined* with similar knowledge from other projects to improve future designs, even if it was not itself set up for this purpose. Likewise, existing research-based generalizations *can be localized* to serve the needs of assessment and improvement for particular projects (and in general they need to be localized to be effective).

Possible recommendations to be made across all topics were also discussed. The final recommendations, based on input from all three meetings and the review of the literature, are included in the next section of the report.

Conclusions and Recommendations

Overview

Effective assessment for informal learning activities needs to observe the following general principles to the extent possible:

- It should take into account the social and cultural context of the communities and institutions supporting and constraining the activities being assessed.
- It should adopt a longitudinal design, following changes and gains from the inception of a project and documenting the processes by which desirable outcomes come about and not the outcomes alone.
- It should adopt a broad definition of knowledge and learning which includes know-how and know-who as well as know-that, and which recognizes that groups and projects as well as individuals learn to do better.
- It should aim to generate information that will be of practical use to individual learners, groups, projects, and organizations not only in identifying outcomes but also in planning for future improvements.

In the introduction to this report, we identified three of the important goals and contributions of this project:

A Model for Good Assessment Practice

The general considerations in the introduction and the specific conclusions and recommendations here define a model of good assessment that extends the definition of learning outcomes beyond discipline-specific knowledge to include valued aspects of social-emotional development, identity, comfort levels, interest and motivation, persistence in the face of challenges and frustrations, effectiveness as members of groups and in using digital technologies, and facility in integrating understanding across activities and domains. The full range of such valued outcomes for individuals is given in more detail in the Introduction.

In addition, this model of good assessment practice, derived from the best current work in the field and the recommendations of our expert panels, requires assessment to consider, to the extent feasible, processes of improvement on the part of groups and teams, whole projects, and participating organizations over a range of timescales from what is learned by individuals in an hour to what may be learned by organizations over a period of years.

Informal learning activities with the key features we have identified can contribute significantly to this comprehensive inventory of valued learning outcomes for individuals, groups, and organizations (see, Introduction and Exhibit 1).

A Toolkit of Assessment Methods and Techniques

In the review of the literature provided in this report, we highlight the variety of methods of documentation and assessment that have proven valuable or are considered promising and worthy of further support and investigation. Among these are broad approaches such as Evidence-Centered Design of assessment and methods derived from it, the use of video documentation and both close analysis and category-based data mining methods of video archives, embedded and unobtrusive integrated assessment in learning games and simulations, longitudinal collections of learner- and group- products, ethnographic observation and fieldnotes, and interviews with participants and organizational stakeholders.

Pointers to the Wider Literature

In addition to the accounts of particular projects reviewed here, we also include in the report a more comprehensive bibliography which focuses on projects and discussions of assessment of informal learning activities that have been influential or deserve to be, many of which were recommended to us by members of our expert panels, and from which we made our selection of the studies to be reviewed. Our bibliography does not focus on the evolving theoretical literature on validity in assessment, though we recognize its relevance and the convergence of some strands in this work with our own conclusions.

General Conclusions

Assessment in Context

The assessment of informal learning activities needs to be specific to the goals of each project and activity, and take account of the history of the project, its supporting organizations and surrounding community. Assessment must also be so organized that it can recognize unanticipated valuable outcomes and processes in progress.

Unit of Analysis and Timescales

The unit of analysis for assessment should be a cluster of related activities across a range of timescales from an instance of an activity; to a sequence of activities that build on one another's content, themes, and skills; to the duration of a working group; to the entire time of an individual's participation in the project; to the full timespan of the project.

Documenting Processes and Outcomes

Documentation of activities needs to support all these scales of assessment for both anticipated and unanticipated outcomes, as well as documenting the over-time processes by which the outcomes are produced. Effective documentation should normally include video records, samples of products produced by participants, interviews with participants in all roles, and fieldnotes made by more than one observer.

Comprehensive View of Outcomes

Valued outcomes include more than just acquired knowledge. Assessment should examine evidence that knowledge is being used (know-how, knowing how to take the next step in an activity) and that this use persists, grows, and cumulates over relatively long periods of time. Relevant knowledge includes knowing how to collaborate with others and who can contribute to meeting a need (know-who). In addition to growth of knowledge, social-emotional development is equally important, including increased persistence in the face of challenges, more effective collaboration with others, increased self-respect, responsibility, initiative, and a sense of agency.

Practical Value of Assessment

Professional quality assessment should provide information and interpretation that is useful to individuals in gauging their changing strengths and weaknesses, for groups to gauge and improve their effectiveness, and for projects and organizations to determine whether goals are being met and to make changes that will better support the learning and development of participants, of groups, and of the project or organization as a whole.

Sustainability and Community Support

Research level assessment should, in addition, aim to provide insight into how and why various features of projects and activities have or have not supported valued learning outcomes and learning-in-progress, for

individuals, groups, and whole projects or organizations. It should also aim to provide a basis for estimating the likelihood that a project is sustainable under various expectable conditions and the likelihood that its successful practices can be adapted to other, specified contexts. Key indices of sustainability include forms and extent of community co-sponsorship and financial and in-kind contributions.

Promising Directions

In the course of the review of the literature and the meetings of our expert panels, the following were identified as promising new directions for informal learning activities, projects, and methods of documentation and assessment, which are deserving of support and further investigation of their effectiveness and usefulness:

- Projects developed and implemented by partnerships between sponsoring organizations (such as universities and informal science education institutions) and local community-based organizations whose leadership and members can represent the perspectives and perceived needs of the community where the project will operate.
- Projects that expand the numbers and kinds of mentors (such as undergraduates, older-age students, and non-education professionals), who can assist in informal learning activities, and which improve their training and experience in doing so.
- Data-mining and machine-learning analysis of logs of interaction with computer programs, video archives, fieldnote archives, and databases of participant products to identify patterns of successful and unsuccessful practices on the part of participants and project organizers.
- The adaptation of the Evidence-Centered Design model to planning more comprehensive documentation and assessment of projects.
- The development of computer-assisted learning games which automatically and unobtrusively document learning progress. These can have a wider range of application if they can be accessed through mobile computing devices (phones, tablets) even if they run on a remote server.
- Ethnographic research in online virtual worlds where new kinds of social learning are taking place. This research can lay the foundation for developing comprehensive documentation and assessment in these new environments.
- The crowd-sourcing of assessment (e.g. the awarding of badges recognizing achievement by peers or senior peers in a project community) can enlist the knowledge and judgment of those closest to an activity. Care should be taken, however, that peer judgments are not superseded by external, standardized criteria.
- The use of agent-based modeling to interrogate and attempt to simulate key processes and practices of participants which contribute to valued outcomes.
- The collection of comprehensive documentation of activity and process in richly instrumented spaces designed for this purpose. Needs to be accompanied by development of tools to analyze the wealth of data collected in newly useful ways.

A Comment on Standardized Testing

Our discussion and general conclusions conspicuously do not recommend the use of standardized testing measures in the assessment of informal learning activities, based on the literature we have reviewed and the opinions in our expert panels.

In the projects we have reviewed standardized measures and test scores play a very small role in assessment, and in the discussions in our expert panel meetings, there was little or no sentiment expressed favoring the relevance or greater use of such measures.

Our view is that the validity of standardized testing measures has not been established for informal learning activities, in large part because these activities do not have fixed, predictable curricular outcomes and because the valued outcomes they do have are either unpredictable or not ones for which valid standard measures exist and which can be used in the informal learning context without being disruptive of other goals and commitments (e.g. to enjoyment, to fostering creative production, to maintaining a free-choice community, etc.). So their use, while not excluded and sometimes of value, should not be considered a norm or ideal of good assessment for informal learning. In fact when they are used, the outcome may say as much about whether the test fits the activity (i.e. about its content and context validity for this particular use) as about whether the activity is producing higher scores on what the test tries to measure.

Policymakers and funders do need assessments that provide a basis for comparison across projects and approaches. The development of such assessments for informal learning activities will require research that can compare and align project-specific assessments across multiple sites and projects over time (see below). It may also be useful to identify the common characteristics of successful informal learning activities across projects and use these as benchmarks when considering new proposals, while still encouraging innovation.

Specific Recommendations

In addition to the General Conclusions above, we have identified the following more specific recommendations to the Foundation and to the research and practice community:

1. Funding for informal learning projects should include set-aside funds for professional assessment, to begin as early in the work of the project as possible, to document learning across the range of outcomes and levels identified in this report, to develop and apply project-specific and activity-specific measures, to modify these as new elements in the project emerge, and to iteratively improve the local assessment model in close interaction with project participants in all roles.

2. Follow-up assessment of the valued consequences of participating in projects one year or more after the end of participation should be undertaken wherever possible.

3. Longitudinal and on-going assessment of projects with multi-year durations is needed, with a focus on change, improvement and sustainability.

4. The development of data-mining and learning-analytics tools which can be applied to a range of data types, including video and fieldnote documentation archives and participant work portfolios, and which can produce results which have practically useful interpretations.

5. It would be valuable to have more research which examines the relationship between play and learning and learning trajectories in activities where playfulness and enjoyment are the dominant mood and motive rather than more exclusively serious approaches to learning for its own sake.

6. The same assessment team should have the opportunity to work with more than one project in order to help develop indices and measures with validity across more than one project, or to document why this strategy may have limited usefulness.

7. Research on the nature and distinctive features of informal learning communities and how they may differ from school-based learning groups is important for improving understanding of social learning in informal activities and for estimating the transferability of successful practices from informal to formal settings.

We wish to thank the MacArthur Foundation for sponsoring this effort and the many generous colleagues who have contributed to deepening our understanding. Effective documentation and assessment of informal learning activities have much to teach us about how and why a wide range of valued outcomes result from such activities. They may also have much to contribute in the future to re-thinking assessment in formal education as well. We hope that this report provides a basis for thoughtful discussion, a useful account of the current state of the art, and a challenge to all of us to do better.

69

References

Anderson, D., and Nashon, S. (2007). Predators of knowledge construction: Interpreting students' metacognition in an amusement park physics program. Science Education, 91(2), 298-320.

Ash, D., Crain, R., Brandt, C., Loomis, M., Wheaton, M., & Bennett, C. (2007). Talk, Tools, and Tensions: Observing biological talk over time. International Journal of Science Education, 29(12), 1581-1602.

Austin, K., Ehrlich, S., Puckett, C., & Singleton, J. (2011). *YouMedia Chicago: Reimagining learning, literacy, and libraries*. Chicago: YouMedia.

Bauman, R. & Briggs, C. (1990). Poetics and performance as critical perspectives on language and social life. *Annual Review of Anthropology, 19*, 59-88.

Behrens, J. T., Mislevy, R. J., Bauer, M., Williamson, D. M., & Levy, R. (2004). Introduction to Evidence Centered Design and Lessons Learned from Its Application in a Global E-Learning Program. International Journal of Testing, 4(4), 7.

Bell, P., Lewenstein, B., Shouse, A., & Feder, M. (Eds.). (2009). *Learning Science in Informal Environments: People, Places and Pursuits.* Washington, DC: National Academy Press.

Bethesda Softworks LLC. *The Elder Scrolls IV: Oblivion, 2006*

Bevan, B. & Michalchik, V. (2012). Out-of-School Time STEM: It's Not What You Think. In B. Bevan, P. Bell, R. Stevens & A. Razfar (Eds.). In *LOST Opportunities: Learning in Out-of-School Time*. New York: Springer.

Blikstein, P., (2011). Using Learning Analytics to Assess Students' Behavior in Open-Ended Programming Tasks. In *Learning Analytics Knowledge Conference.* Banff, Canada

Brandes, U., and Erlebach, T., (2005). *Network analysis: Methodological foundations.* Berlin, Heidelberg: Springer-Velag.

Burghardt, M. D., (1999). *Introduction to engineering design and problem solving.* Boston, MA: WCB/McGraw-Hill.

Chen, M. (2009). Communication, Coordination, and Camaraderie in World of Warcraft. *Games and Culture 4*: 47-73.

Cole, M. and The Distributed Literacy Consortium. (2006). The Fifth Dimension: An after-school program built on diversity. Russell Sage Foundation.

Dahotre, A, Zhang, Y, and Scaffidi, C. 2010. A qualitative study of animation programming in the wild. Symp. on Empirical Software Engineering and Measurement, 1-10.

Downing-Wilson, (2006/2007). Revealing shifts in attitude among undergraduates participating in academic service learning programs. *Operant Subjectivity*, 2006/2007 (October/January), 30(1/2): 23 -51.

Downing-Wilson, D., Lecusay, R., Rosero, I., Cole, M. (2012). A Cultural-Historical Approach toUniversity/Community Collaborative Interventions. In J. Valsiner (ed.) *Encyclopedia of Cultural Psychology* (pp. 885-898) . New York, NY: Oxford University Press.

Dym, C. L., & Little, P. (2000). *Engineering design: A project-based introduction.* New York, NY: Wiley.

Eccles, J., (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly, 18,* 585-609.

Eisenhart, M., and Edwards, L. (2004). Red-eared sliders and neighborhood dogs: Creating third spaces to support ethnic girls' interests in technological and sci- entific expertise. Children, Youth and Environments, 14(2), 156-177. Available: http://www.colorado.edu/journals/cye/ [accessed October 2008].

Fadigan, K.A., and Hammrich, P.L. (2004). A longitudinal study of the educational and career trajectories of female participants of an urban informal science education program. Journal of Research on Science Teaching, 41(8), 835-860.

Fields, D. and Kafai, Y. (2007). Tracing Insider Knowledge Across Time And Spaces: A Comparative Ethnography in a Teen Online Game World. Proceedings Conference on Computer-supported Collaborative Learning. Pp. 230-239.

Gallagher, J. J., & Tobin, K. G. (1991). Reporting interpretive research. In J. Gallagher (Ed.), *Interpretive research in science education* (pp. 85-95). National Association of Research in Science Teaching Monograph No. 4. Manhattan, KS: NARST.

Gee, J. (2003). What Video Games have to Teach Us about Learning and Literacy. New York: Palgrave Macmillan.

Gutiérrez, K., Rymes. B. & Larson, J. (1995). Script, counterscript and underlife in the classroom: James Brown versus Brown v. The Board of Education. *Howard Educational Review, 65(3): 445-47*

Glaser, B. G., & Stauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research. Chicago, IL: Aldine.*

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience.* New York, NY: Harper & Row.

Gutwill, J. P., & Allen, S. (2010). Facilitating family group inquiry at science museum exhibits. Science Education, 94(4), 710-742. doi:10.1002/sce.20387

Harvard Family Research Project. "After School Programs in the 21st Century: Their Potential and What It Takes to Achieve It." No. 10. Issues and Opportunities in Out-of-School Time Evaluation. Cambridge, MA: Harvard Family Research Project, February 2008.

Hines, C. (2000). *Virtual ethnography*. Thousand Oaks, CA: Sage.

Hull, G., & James, M. (2007). Geographies of hope: A study of urban landscapes and a university-community collaborative. In Peggy O'Neill (Ed.), Blurring boundaries: Developing writers, researchers, and teachers: A tribute to William L. Smith (pp. 250-89). Cresskill, NJ: Hampton Press.

Hull, G. A., & Katz, M. L. (2006). Crafting an agentive self: Case studies of digital storytelling. Research in the Teaching of English, 43-81.

Hutchins, E. (1995). Cognition in the wild. Cambridge, MA: MIT Press. King, A. (1990). Facilitating elaborative learning in the classroom through reciprocal questioning. *American Educational Research Journal, 27*, 664–687.

Hsi, S., 2003, A study of user experiences mediated by nomadic web content in a museum. *Journal of Computer Assisted Learning. 19*, 308-319

Kafai, Y., Peppler, K., Alavez, M., & Ruvalcaba, O. (2006). *Seeds of a computer culture: An archival analysis of programming artifacts from a community technology center.* In Barab, S., Hay, K., Hickey, D. (eds.). Proceedings of the Seventh International Conference of the Learning Sciences (pp. 942-943).

Kafai, Y.B., Peppler, K., & Chapman, R. (Eds.) (2009). *The Computer Clubhouse: Creativity and Constructionism in Youth Communities*. New York, NY: Teachers College Press.

Kafai, Y. B., Peppler, K. A., & Chiu, G. M. (2007). High Tech Programmers in Low-Income Communities: Creating a Computer Culture in a Community Technology Center. Communities and Technologies 2007, 545-563.

Kuhn, A., McNally, B., Schmoll, S., Cahill, C., Lo, W., Quintana, C., and Delen, I. (2012). *How students find, evaluate, and utilize peer-collected annotated multimedia data in science inquiry with Zydeco*. In Proceedings of the 2012 annual conference on Human factors in computing systems (CHI '12). ACM, New York, NY, USA, 3061-3070.

Leander, K., & Mckim, K. (2003). Tracing the everyday 'sitings' of adolescents on the Internet: A strategic adaptation of ethnography across online and offline spaces. *Education, Communication & Information, 3 (2): 211-240*

Mahoney, J., Larson, R., Eccles, J. S. (Eds.) (2005). Organized activities as contexts of development: Extracurricular activities, after-school, and community programs. Hillsdale, NJ: Erlbaum.

Maloney, J. H., Peppler, K., Kafai, Y., Resnick, M., & Rusk, N. (2008). Programming by choice: urban youth learning programming with scratch. ACM SIGCSE Bulletin, 40(1), 367-371.

Maxwell, J. A., (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.

Merriam, S. B., (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey-Bass.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). A brief introduction to evidence- centered design (CSE Report 632). CA: Center for Research on Evaluation, Standards, and Student Testing. (ERIC Document Reproduction Service No. ED483399)

Mislevy, R. J., and Haertel, G. D. (2007) Implication of Evidence-Centered Design for Educational Testing, *Educational Measurement: Issues & Practive*. http://www.em2007.mpg.de/files/MislevyHaertel.pdf

Nacu, N., Pinkard, K., Larson, K., and Schmidt, R., (2012) *IPE: Challenge-based Learning and the PLAYground: What is challenge-based learning and how can we use* an online platform to explore it*?* Proceedings of the 2012 Digital Media & Learning Conference: Beyond Educational Technology, San Francisco, CA

National Research Council. (2000). How people learn: Brain, mind, experience, and school.Committee on Developments in the Science of Learning, J.D. Bransford, A.L. Brown, and R.R. Cocking (Eds.), and Committee on Learning Research and Educational Practice, M.S. Donovan, J.D. Bransford, and J.W. Pellegrino (Eds.). Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Oakes, J. (1992). Can tracking research inform practice? Technical, normative, and political considerations. *Educational Researcher, 21(4)*, 12-21.

Quintana, C., Cahill, C., Kuhn, A., Schmoll, S., and Pompe, A. (2010). Zydeco: using mobile and web technologies to support seamless inquiry between museum and school contexts. *In Proceedings of the 9th International Conference on Interaction Design and Children*. (IDC '10). ACM, New York, NY, USA, 174-177

Rahm, J. (2002). Emergent learning opportunities in an inner-city youth gardening program. *Journal of Research in Science Teaching, 39*(2), 164-184.

Resnick, M., Rusk, N., and Cooke, S. (1998). The Computer Clubhouse: Technological Fluency in the Inner City. In Schon, D., Sanyal, B., and Mitchell, W. (eds.), *High Technology and Low-Income Communities*, pp. 266-286. Cambridge: MIT Press.

Salen, K., Torres, R., Wolozin, L., Rufo-Tepper, R., & Shapiro, A. (2010). Quest to Learn: Developing the School for Digital Kids. The MIT Press.

Scaffidi, C., & Chambers, C. (2011). Skill progression demonstrated by users in the Scratch animation environment. International Journal of Human-Computer Interaction

Shaffer, D.W. (2005). Epistemography and the Participant Structures of a Professional Practicum: A story behind the story of Journalism 828 (WCER Working Paper No. 2005-8): University of Wisconsin-Madison, Wisconsin Center for Education.

Shaffer, D.W. (2006). How computer games help children learn. New York: Palgrave Macmillan.

Shaffer, D.W. (2007). In praise of epistemology. In Brett E. Shelton and David A. Wiley (Eds.), The design and use of simulation computer games in education (pp. 7-27). Rotterdam, The Netherlands: Sense Publishers

Shaffer, D. W., & Gee, J. P. (2012). The Right Kind of GATE: Computer games and the future of assessment.. In M. Mayrath, J. Clarke-Midura, & D. H. Robinson (Eds.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research*. Charlotte, NC: Information Age Publishing.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., et al. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning. International Journal of Learning, 1(2), 33-53.

Shute, V. J. (2007). Tensions, trends, tools, and technologies: Time for an educational sea change. In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 139-187). New York, NY: Lawrence Erlbaum Associates, Taylor & Francis Group.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Shute, V. J., Hansen, E. G., & Almond, R. G. (2008). You can't fatten a hog by weighing it—or can you? Evaluating an assessment for learning system called ACED. *International Journal of Artificial Intelligence and Education, 18*(4), 289-316.

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. Computer games and instruction. Charlotte, NC: Information Age Publishers.

Shute, V. J., & Torres, R. (2012). Where streams converge: Using evidence-centered design to assess Quest to Learn. Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research. Charlotte, NC: Information Age Publishing.

Shute, V. J., Ventura, M., Bauer, M. I., & Zapata-Rivera, D. (2009). Melding the pow- er of serious games and embedded assessment to monitor and foster learn- ing: Flow and grow. In U. Ritterfeld, M. Cody, & P. Vorderer (Eds.), Serious games: Mechanisms and effects (pp. 295–321). Mahwah, NJ: Routledge, Taylor and Francis.

Spradley, J. P., (1979) *The Ethnographic Interview*. New York: Holt. Rienhart and Winston.

Spradley, J. P., (1980) *Participant Observation*. New York: Holt. Rienhart and Winston.

Squire, K. (2006). From content to context: Videogames as designed experience. Educational Researcher, 35(8), 19-29.

State, R. E., (1995). *The art of case study research.* Thousand Oaks, CA: Sage.

Steinkuehler, C. (2004). Learning in massively multiplayer online games. Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.). Proceedings of the Sixth International Conference of the Learning Sciences, (521–528). Mahwah, NJ: Erlbaum.

Steinkuehler, C. A. (2005). *Cognition and learning in massively multiplayer online games: A critical approach.* University of Wisconsin-Madison, Madison, WI.

Steinkuehler, C. A. (2006). Why game (culture) studies now? Games and Culture, 1(1), 1-6.

Steinkuehler, C. (2008). Cognition and literacy in massively multiplayer online games. In D. Leu, J. Coiro, C. Lankshear, & K. Knobel (Eds.). Handbook of Research on New Literacies, (pp. 611- 634). Mahwah NJ: Erlbaum.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research* (2nd ed.). Thousand Oaks, CA: Sage

Svarovsky, G. N., & Shaffer, D. W. (2006a). Design meetings and design notebooks as tools for reflection in the engineering design course. Paper presented at the Paper presented at the 36th ASEE/IEEE Frontiers in Education Conference, San Diego, CA.

Svarovsky, G. N., & Shaffer, D. W. (2006b). SodaConstructing an understanding of physics: Technology-based engineering activities for middle school students. Paper presented at the Paper presented at the 36th ASEE/IEEE Frontiers in Education Conference, San Diego, CA.

Urciuoli, B. (1995). The indexical structure of visibility. In B. Farnell (Ed.), *Human action signs in cultural context: The visible and the invisible in movement and dance* (pp.189-215). Metuchen, NJ: The Scarecrow Press.

Van Maanen, J. (1988). *Tales of the field*. Chicago, IL: Chicago University Press

Wilensky, U. (1999). GasLab–an Extensible Modeling Toolkit for Exploring Micro- and Macro- Views of Gases. In Roberts, N., Feurzeig, W. & Hunter, B. (Eds.) *Computer Modeling and Simulation in Science Education*. Berlin: Springer Verlag.

Appendix A: Expert Meeting Participants

San Francisco, June 10-11, 2011
Mike Cole (UC, San Diego)
Bronwyn Bevan (Exploratorium)
Shelley Goldman (Stanford)
Louis Gomez (UC, Los Angeles)
Kris Gutierrez (UC, Los Angeles)
Robert Lecusay (UC, San Diego)
Jay Lemke (UC, San Diego)
Barbara Means (SRI)
Ray McDermott (Stanford)
Vera Michalchik (SRI)
Bonnie Nardi (UC, Irvine)

Northwestern, October 21-22, 2011
Sasha Barab (Arizona State University)
Mike Cole (UC, San Diego)
Kevin Leander (Vanderbilt)
Robert Lecusay (UC, San Diego)
Jay Lemke (UC, San Diego)
Vera Michalchik (SRI)
Nichole Pinkard (DePaul University)
Jan Plass (NYU)
Laura Salus (Northwestern)
Reed Stevens (Northwestern)
Carla Seal Wanner (Flickerlab)
Stanton Wortham (U Penn)

San Diego, February 3-4, 2012
Phil Bell (University of Washington)
Paulo Blikstein (Stanford)
Mike Cole (UC, San Diego)
Jim Gee (Arizona State University)
Mimi Ito (UC, Irvine)
Robert Lecusay (UC, San Diego)
Jay Lemke (UC, San Diego)
Vera Michalchik (SRI)

Bill Penuel (University of Colorado, Boulder)
Nichole Pinkard (DePaul University)
Reed Stevens (Northwestern)
Jennifer Vadeboncoeur (University of British Columbia)

Appendix B: Bibliography

Afterschool Alliance (2011). A Summary of Formal Evaluations of Afterschool Programs' Impact on Academics, Behavior, Safety and Family Life

Afterschool Alliance. (2008). 21st century learning centers providing supports to com- munities nationwide. Available: http://www.afterschoolalliance.org/researchFact Sheets.cfm [accessed October 2008].

Akilli, G. K. (2007). Games and Simulations: A new approach in Education? In Gibson, D., Aldrich, C. and Prensky, M. (eds.), Games and Simulations in Online Learning: Research and Development Frameworks (pp.1-20). Hershey PA: Information Science Publishing.

Aldrich, C. (2004). Simulations and the future of learning. San Francisco. CA. John Wiley and Sons.

Allen, S. (2002). Looking for learning in visitor talk: A methodological exploration. In G. Leinhardt, K. Crowley, and K. Knutson (Eds.), Learning conversations in museums (pp. 259-303). Mahwah, NJ: Lawrence Erlbaum Associates.

Allen, S., & Gutwill, J. (2004). Designing With Multiple Interactives: Five Common Pitfalls. Curator: The Museum Journal, 47(2), 199-212.

Anderson, D., and Nashon, S. (2007). Predators of knowledge construction: Interpreting students' metacognition in an amusement park physics program. Science Education, 91(2), 298-320.

Anderson, D., Lucas, K. B., & Ginns, I. S. (2003). Theoretical perspectives on learning in an informal setting. Journal of Research in Science Teaching, 40(2), 177-199.

Ang, C. S. and Zaphiris, P. (2008). Social learning in MMOGs: An Activity Theoretical Perspective. Interactive Technology and Smart Education 5: 84-102.

Ansbacher, T. (1997). If technology is the answer, what was the question? Technology and experience-based learning. Hand to Hand, 11(3), 3-6.

Ash, D. (2003). Dialogic inquiry in life science conversations of family groups in a museum. Journal of Research in Science Teaching, 40(2), 138-162.

Ash, D., Crain, R., Brandt, C., Loomis, M., Wheaton, M., & Bennett, C. (2007). Talk, Tools, and Tensions: Observing biological talk over time. International Journal of Science Education, 29(12), 1581-1602.

Astor-Jack, T., Whaley, K.K., Dierking, L.D., Perry, D., and Garibay, C. (2007). Understanding the complexities of socially mediated learning. In J.H. Falk, L.D. Dierking, and S. Foutz (Eds.), In principle, in practice: Museums as learning institutions. Walnut Creek, CA: AltaMira Press.

Augustin, T., Hockemeyer, C., Kickmeier-Rust, M., & Albert, D. (2011). Individualized skill assessment in digital learning games: Basic definitions and mathematical formalism. Learning Technologies, IEEE Transactions on, (99), 1-1.

Austin, K., Ehrlich, S., Puckett, C., & Singleton, J. (2011). *YouMedia Chicago: Reimagining learning, literacy, and libraries*. Chicago: YouMedia.

Azevedo, F.S. (2004). Serious play: A comparative study of learning and engagement in hobby practices. Berkeley: University of California Press.

Azevedo, F.S. (2006). Personal excursions: Investigating the dynamics of student engagement. International Journal of Computers for Mathematical Learning, 11(1), 57-98.

Bainbridge, W. (2007). The Scientific Research Potential of Virtual Worlds. Science 317: 472-476.

Banks, J.A., Au, K.H., Ball, A.F., Bell, P., Gordon, E.W., Gutiérrez, K., Heath, S.B., Lee, C.D., Lee, Y., Mahiri, J., Nasir, N.S., Valdes, G., and Zhou, M. (2007). Learning in and out of school in diverse environments: Lifelong, life-wide, life-deep. Seattle: Center for Multicultural Education, University of Washington.

Barab, S., Thomas, M. Dodge, T., Carteaux, R. and Tuzun, H. (2005). Making learning fun: Quest Atlantis, agame without guns. Educational Technology Research and Development, 53(1): 86-107.

Barab, S., Zuiker, S. Warren, S., Hickey, D., Ingram-Goble, A., Kwon, E-J., Kouper, I. and Herring, S. (2007). Situationally Embodied Curriculum: Relating Formalisms and Contexts. WileyInterScience Online Periodicals.

Barab, S.A., Dodge, T., Ingram-Goble, A., Peppler, K., Pettyjohn, P., Volk, C.,& Solomou, M. (2010). Pedagogical dramas and transformational play: Narratively-rich games for learning. Mind, Culture, and Activity 17(3): 235–264.

Barab, S.A., Gresalfi, M.S., & Ingram-Goble, A. (2010). Transformational play: Using games to position person, content, and context. Educational Researcher, 39(7), 525-536.

Barab, S.A., Gresalfi, M.S., Dodge, T., & Ingram-Goble, A. (2010). Narratizing Disciplines andDisciplinizing Narratives: Games as 21st Century Curriculum. International Journal for Gaming and Computer-Mediated Simulations, 2(1), 17-30.

Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. Human Development, 49(4), 153-224.

Bartko, W. T. (2005). The ABCs of engagement in out□of□school□time programs. New Directions for Youth Development, 2005(105), 109-120.

Beals, D.E. (1993). Explanatory talk in low-income families' mealtime. Preschoolers' questions and parents' explanations: Causal thinking in everyday parent-child activity. Hispanic Journal of Behavioral Sciences, 19(1), 3-33.

Beaumont, L. (2005). Summative evaluation of wild reef-sharks at Shedd. Report for the John G. Shedd Aquarium. Available: http://www.informalscience.com/ download/case_studies/report_133.doc [accessed October 2008].

Behrens, J. T., Mislevy, R. J., Bauer, M., Williamson, D. M., & Levy, R. (2004). Introduction to Evidence Centered Design and Lessons Learned from Its Application in a Global E-Learning Program. International Journal of Testing, 4(4), 7.

Bell, P., Bricker, L.A., Lee, T.R., Reeve, S., and Zimmerman, H.T. (2006). Understanding the cultural foundations of children's biological knowledge: Insights from everyday cognition research. In S.A. Barab, K.E. Hay, and D. Hickey (Eds.), Pro- ceedings of the seventh international conference of the learning sciences (ICLS) (pp. 1029-1035). Mahwah, NJ: Lawrence Erlbaum Associates.

Bell, P., Zimmerman, H.T., Bricker, L.A., and Lee, T.R. (no date). The everyday cultural foundations of children's biological understanding in an urban, high- poverty community. Everyday Science and Technology Group, University of Washington.

Bevan, B. & Michalchik, V. (2012). Out-of-School Time STEM: It's Not What You Think. In B. Bevan, P. Bell, R. Stevens & A. Razfar (Eds.). In *LOST Opportunities: Learning in Out-of-School Time*. New York: Springer.

Bevan et al. (2010). Making Science Matter: Collaborations Between Informal Science Education Organizations and Schools. Washington, DC: Center for Advancement of Informal Science Education.

Bieber, A. E., Marchese, P., & Engelberg, D. (2005). The laser academy: An after-school program to promote interest in technology careers. Journal of Science Education and Technology, 14(1), 135-142.

Birmingham, J., Pechman, E., Russell, C. and Mielke, M. Shared Features of High-Performing After-School Programs: A Follow-Up to the TASC Evaluation. Washington, DC: Policy Studies Associates, November 2005.

Black, A. R., Doolittle, F. C., Zhu, P., Unterman, R., & Grossman, J. B. (2008). The evaluation of enhanced academic instruction in after-school programs: Findings after the first year of implementation.

Blackburn, M. V. (2002). Disrupting the (hetero) normative: Exploring literacy performances and identity work with queer youth. Journal of Adolescent & Adult Literacy, 46(4), 312-324.

Blanton, W. E., Moorman, G. B., Hayes, B. A., & Warner, M. L. (1997). Effects of partici- pation in the Fifth Dimension on far transfer.Journal ofEducational Computing Research, 16, 371-396.

Blanton, W. E., Moorman, G. B., Hayes, B. A., & Warner, M. L. (1997). Effects of participation in the Fifth Dimension on far transfer. Journal of Educational Computing Research, 16(4), 371-396.

Blanton, W., Mayer, R., Gallego, M., McNamee, G., Shustack, M. (2006) The quantitative effects of Fifth Dimension participation on children's cognitive and academic skills. In M. Cole & the Distributed Literacy Consortium (Eds.), The fifth dimension: An after-school program built on diversity (pp. 107–128), New York: Russell Sage Foundation.

Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. Studies in continuing education, 22(2), 151-167.

Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. Assessment & Evaluation in Higher Education, 24(4), 413-426.

Boudria, T. J. (2002). Implementing a project-based technology program for high school women. Community College Journal of Research &Practice, 26(9), 709-722.

Bouillion Diaz, L. (2009) Creating Opportunities for Ubiquitous Learning with Geospatial Technologies: Negotiating Roles at the Borders of Youth and Adult Practice, In Cope, B. and Kalantzis, M. (Eds) Ubiquitous Learning. Champaign, IL. University of Illinois Press.

Bourgonjon, J., Rutten, K., Soetaert, R., & Valcke, M. (2011). From Counter-Strike to Counter-Statement: using Burke's pentad as a tool for analysing video games. Digital Creativity, 22(2), 91-102.

Bremme, D., Blanton, W., Gallego, M., Moll, L. C., Rueda, R., & Va´squez, O. (2006). The dynamics of change in children's learning. In M. Cole & the Distributed Literacy Consortium (Eds.), The fifth dimension: An after-school program built on diversity (pp. 107–128),New York: Russell Sage Foundation.

Brody, M., Bangert, A., Dillon, J. (2009). Assessing Learning in Informal Science Contexts. Commissioned paper by the National Research Council for Science Learning in Informal Environments Committee.

Bronte-Tinkew, J., Moore, K. and Shwalb, R. Measuring Outcomes for Children and Youth in Out-Of-School Time Programs: Moving Beyond Measuring Academics. Washington, DC: Child Trends, October 2006.

Brosi, E. (2011) Measurement Tools for Evaluating Out-of-School Time Programs: An Evaluation Resource

Brown & Ferrara (1985) Diagnosing zones of proximal development

Brown, B.A. (2006). "It isn't no slang that can be said about this stuff": Language, identity, and appropriating science discourse. Journal of Research in Science Teaching, 43(1), 96-126.

Bruce, B. C., Bruce, S., Conrad, R. L., & Huang, H. J. (1997). University science students as curriculum planners, teachers, and role models in elementary school classrooms.

Buckingham, D. (2008). Youth, identity, and digital media. The MIT Press.

Buckingham, D., & Scanlon, M. (2003).Education, Entertainment and Learning in the Home. Milton Keynes: Open University Press.

Bull, G., Thompson, A., Searson, M., Garofalo, J., Park, J., Young, C., & Lee, J. (2008). Connecting informal and formal learning: Experiences in the age of participatory media. Contemporary Issues in Technology and Teacher Education, 8(2), 100-107.

Burke, A., & Hammett, R. F. (2009). Assessing new literacies: Perspectives from the classroom (Vol. 34). Peter Lang Pub Inc.

Calabrese Barton, A. (1998b). Reframing "science for all" through the politics of poverty. Educational Policy, 12, 525-541.

Callanan, M.A., and Oakes, L. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. Cognitive Development, 7, 213-233.

Campbell, P. (2008, March). Evaluating youth and community programs: In the new ISE framework. In A. Friedman (Ed.), Framework for evaluating impacts of informal science education projects (pp. 69-75). Available: http://insci.org/docs/ Eval_Framework.pdf [accessed October 2008].

Carnegie Council on Adolescent Development (1992). A matter of time: Risk and opportunity in the nonschool hours. Carnegie Council Monograph. http://www.carnegie.org/ccadpubs.htm

Catterall, J. (2002), Involvement in the Arts and Success in Secondary School. In R. Deasy (Ed.), Critical links: Learning in the arts and student achievement and social development. Washington, DC: AEP.

Chen, M. (2009). Communication, Coordination, and Camaraderie in World of Warcraft. Games and Culture 4: 47-73.

Choontanom, T. and Nardi, B. (2010). Theorycrafting: The Art and Science of Using Numbers to Interpret the World. Forthcoming in Games, Learning and Society. C. Steinkuehler, ed. London: Cambridge University Press.

Chung, A. and Hillsman, E. “Evaluating After-School Programs.” The School Administrator. Arlington, VA: American Association of School Administrators, May 2005.

Clarke, J., & Dede, C. (2009). Design for Scalability: A Case Study of the River City Curriculum. Journal of Science Education and Technology, 18(4), 353-365.

Cobern, W.W., and Aikenhead, G.S. (1998). Cultural aspects of learning science. In B. Fraser and K. Tobin (Eds.), International handbook of science education (Part One, pp. 39-52). Dordrecht, Netherlands: Kluwer Academic.

Colardyn, D., & Bjornavold, J. (2004). Validation of Formal, Non□Formal and Informal Learning: policy and practices in EU Member States1. European Journal of Education, 39(1), 69-89.

Cole, M. and The Distributed Literacy Consortium. (2006). The Fifth Dimension: An after-school program built on diversity. Russell Sage Foundation.

Coleman, J. (1971). Learning through games. In E. Avedon and B. Sutton-Smith (Eds). The study of games. New York and London. John Wiley, pp. 322-329.

Colley, H., Hodkinson, P., & Malcolm, J. (2002). Non-formal learning: mapping the conceptual terrain. Consultation report, Leeds: University of Leeds Lifelong Learning Institute. Also available in the informal education archives: http://www. infed. org/archives/etexts/colley_informal_learning. htm.

Colley, H., Hodkinson, P., & Malcolm, J. (2003). Informality and formality in learning: a report for the Learning and Skills Research Centre. Learning and Skills Research Centre London.

Connolly, T.M., Boyle, E.A., Stansfield, M.H., and Hainey, T. (2007) "A Survey of Students' Computer Games Playing Habits", Journal of Advanced Technology for Learning, Issue 4.

Connolly, T.M., Stansfield, M.H. and Hainey, T. (2009) "Towards the Development of a Games-based Learning Evaluation Framework". Games-based Learning Advancement for Multisensory Human Computer Interfaces: Techniques and Effective Practices (Eds: Connolly, T.M., Stansfield, M.H. and Boyle, E.). Idea-Group Publishing: Hershey.

COSMOS Corporation. (1998). A report on the evaluation of the National Science Foundation’s informal science education program. Washington, DC: National Science Foundation. Available: http://www.nsf.gov/pubs/1998/nsf9865/nsf9865. htm [accessed October 2008].

Costa, V.B. (1995). When science is “another world”: Relationships between worlds of family, friends, school, and science. Science Education, 79, 313-333.

Crane, V., Nicholson, H.J., Chen, M., and Bitgood, S. (1994). Informal science learning: What the research says about television, science museums and community-based projects. Dedham, MA: Research Communications.

Crowley, K. D., Schunn, C. D., & Okada, T. (2001). Designing for science: Implications from everyday, classroom, and professional settings. Lawrence Erlbaum.

Crowley, K., and Galco, J. (2001). Everyday activity and the development of scientific thinking. In K. Crowley, C.D. Schunn, and T. Okada (Eds.), Designing for science: Implications from everyday, classroom, and professional settings (pp. 123-156). Mahwah, NJ: Lawrence Erlbaum Associates.

Crowley, K., and Galco, J. (2001). Family conversations and the emergence of scientific literacy. In K. Crowley, C. Schunn, and T. Okada. (Eds.), Designing for science: Implications from everyday, classroom, and professional science (pp. 393-413). Mahwah, NJ: Lawrence Erlbaum Associates.

Curtis, D., & Lawson, M. (2002). Computer adventure games as problem-solving environments. International Education Journal, 3(4), 43-56.

Dahotre, A, Zhang, Y, and Scaffidi, C. 2010. A qualitative study of animation programming in the wild. Symp. on Empirical Software Engineering and Measurement, 1-10.

Dancu, T. (2006). Comparing three methods for measuring children's engagement with exhibits: Observations, caregiver interviews, and child interviews. Poster presented at 2006 Annual Meeting of the Visitor Studies Association, Grand Rapids, MI.

Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. Teaching and teacher education, 16(5-6), 523-545.

Davidson, A.L. (1999). Negotiating social differences: Youths' assessments of educators' strategies. Urban Education, 34(3), 338-369.

Davidson, C. N., & Goldberg, D. T. (2009). The future of learning institutions in a digital age. The MIT Press.

Davidson, C. N., & Goldberg, D. T. (2010). The future of thinking: Learning institutions in a digital age. The MIT Press.

de Freitas, S. & Griffiths, M. D. (2008). The convergence of gaming practices with other media forms: what potential for learning? A review of the literature. Learning, Media Technology, 33(1).

de Freitas, S. & Griffiths, M. D. (2009) Massively multiplayer online role-play games to support learning and training communities. In R. Ferdig (Ed.) Handbook of Research on effective electronic gaming in education. Idea Group Publishing.

de Freitas, S., & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? Computers & Education, 46(3), 249-264.

de Freitas, S., Rebolledo-Mendez, G., Liarokapis, F. Magoulas, G. Poulovassilis A. (2010). Learning as immersive experiences: using the four dimensional framework for designing and evaluating immersive learning experiences in a virtual world. British Journal of Educational Technology, 41(1): 69-85.

Dede, C. (2006) Scaling Up: Evolving Innovations beyond Ideal Settings to Challenging Contexts of Practice. In R.K. Sawyer (Ed.), Cambridge Handbook of the Learning Sciences. Cambridge, England: Cambridge University Press.

Dede, C., Rockman, S., & Knox, A. (2007). Lessons learned from studying how innovations can achieve scale. Threshold 5, 1, 4-10.

Defining a Model for Collaborative Research at the Intersection of Learning, Reform Implementation, and Design

Derry, S. (2007). Guidelines for video research in education: Recommendations from an expert panel. Chicago, IL: University of Chicago, Data Research and Development Center.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., Hall, R., et al. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. Journal of the Learning Sciences, 19(1), 3-53.

Diamond, J. (1999). *Practical evaluation guide: Tools for museums and other educational settings*. Walnut Creek, CA: AltaMira Press.

Dickey, M.D. (2006). Girl gamers: The controversy and relevance of female-oriented design for instructional design. *British Journal of Educational Technology*. 37(5). 785–793.

Dierking, L. D., & Falk, J. H. (2003). Optimizing out□of□school time: The role of free□choice learning. New Directions for Youth Development, 2003(97), 75-88.

Dierking, L.D., Adelman, L.M., Ogden, J., Lehnhardt, K., Miller, L., and Mellen, J.D. (2004). Using a behavior changes model to document the impact of visits to Disney's Animal Kingdom: A study investigating intended conservation action. Curator, 47(3), 322-343.

Dierking, Lynn D, & Falk, John H. (1994). Family behavior and learning in informal science settings: A review of the research. Science Education, 78(1), 57-72. doi:10.1002/sce.3730780104

Diversi, M., & Mecham, C. (2005). Latino (a) students and Caucasian mentors in a rural after□school program: Towards empowering adult–youth relationships. *Journal of Community Psychology, 33*(1), 31-40.

Downing-Wilson, (2006/2007). Revealing shifts in attitude among undergraduates participating in academic service learning programs. *Operant Subjectivity*, 2006/2007 (October/January), 30(1/2): 23 -51.

Durlak, J. and Weissburg, R. The Impact of After-school Programs That Promote Personal and Social Skills. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning, January 2007.

Easton, L.B. *Annotated Bibliography on Learning*. Denver, CO: Education Commission of the States, September 2005.

Education week special issue on informal science learning

Egan□Robertson, A. (1998). Learning about culture, language, and power: Understanding relationships among personhood, literacy practices, and intertextuality. Journal of Literacy Research, 30(4), 449-487.

Eisenhart, M., and Edwards, L. (2004). Red-eared sliders and neighborhood dogs: Creating third spaces to support ethnic girls' interests in technological and sci- entific expertise. Children, Youth and Environments, 14(2), 156-177. Available: http://www.colorado.edu/journals/cye/ [accessed October 2008].

Ekman, P., and Rosenberg, E. (Eds.). (2005). What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system. New York: Oxford University Press.

Ellenbogen, K.M. (2002). Museums in family life: An ethnographic case study. In G. Leinhardt, K. Crowley, and K. Knutson (Eds.), Learning conversations in muse- ums. Mahwah, NJ: Lawrence Erlbaum Associates.

Ellenbogen, K.M., Luke, J.J., and Dierking, L.D. (2004). Family learning research in museums: An emerging disciplinary matrix? Available: http://www3.interscience. wiley.com/cgi-bin/fulltext/109062559/PDFSTART [accessed March 2009].

Eraut, M. (2000). Non□formal learning and tacit knowledge in professional work. British Journal of Educational Psychology, 70(1), 113-136.

Erbstein, N. and Bookmyer, J. Afterschool Evaluation Practices in California: Site-level Perspectives. Davis, CA: Center for Community School Partnerships, November 2005.

Fadigan, K.A., and Hammrich, P.L. (2004). A longitudinal study of the educational and career trajectories of female participants of an urban informal science education program. Journal of Research on Science Teaching, 41(8), 835-860.

Fadjo, C.L., Lu, M.T. & Black, J.B. (2009). Instructional Embodiment and Video Game Programming in an After School Program. In G. Siemens & C. Fulford (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009 (pp. 4041-4046). Chesapeake, VA: AACE.

Falk, J.H. (2001). Free-choice science education: How we learn science outside of school. New York: Teachers College Press.

Falk, J.H. and Adleman, L.M. (2003). Investigating the impact of prior knowledge and interest on aquarium visitor learning. Journal of Research in Science Teaching, 40(2), 163-76.

Falk, J.H., and Storksdieck, M. (2005). Using the “contextual model of learning” to understand visitor learning from a science center exhibition. Science Education, 89, 744-778.

Falk, J.H., Moussouri, T., and Coulson, D. (1998). The effect of visitors’ agendas on museum learning, Curator, 41(2), 107-120.

Falk, John Howard, & Dierking, Lynn Diane. (1992). The museum experience. Howells House.

Falk, John Howard, & Dierking, Lynn Diane. (2000). Learning from museums: visitor experiences and the making of meaning. Rowman & Littlefield.

Fancsali, C. (2002). What we know about girls, STEM, and afterschool programs. Prepared for Education Equity Concepts. New York: Academy for Educational Development.

Fashola, O. S., & (US), E. R. I. C. (1998). Review of extended-day and after-school programs and their effectiveness. Center for Research on the Education of Students Placed At Risk, Johns Hopkins University & Howard University.

Feldon, D.F., and Gilmore, J. (2006). Patterns in children’s online behavior and scientific problem-solving: A large-N microgenetic study. In G. Clarebout and J. Elen (Eds.), Advances in studying and designing (computer-based) powerful learning environments (pp. 117-125). Rotterdam, The Netherlands: Sense.

Fender, J.G, and Crowley, K. (2007). How parent explanation changes what children learn from everyday scientific thinking. Journal of Applied Developmental Psy- chology, 28(3), 189-210.

Fenichel, M., & Schweingruber, H. A. (2010). Surrounded by science: Learning science in informal environments. Natl Academy Pr.

Ferreira, M. (2001). The effect of an after-school program addressing the gender and minority achievement gaps in science, mathematics, and engineering. Arlington, VA: Educational Research Spectrum, Educational Research Services.

Fields, D. and Kafai, Y. (2007). Tracing Insider Knowledge Across Time And Spaces: A Comparative Ethnography in a Teen Online Game World. Proceedings Conference on Computer-supported Collaborative Learning. Pp. 230-239.

Foehr, U. G., Roberts, D. F., & Rideout, V. J. (n.d.). Generation M2: Media in the Lives of 8- to 18-Year-Olds, (8010), 79.

Frake, C. O. (1983). Notes toward a cultural analysis of formal. Text-Interdisciplinary Journal for the Study of Discourse, 3(3), 299-304.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59.

Freeman, J., Dorph, R., & Chi, B. S. (2009). Strengthening After School STEM Staff Development. A final report to the Noyce Foundation. Coalition for Science After School. Lawrence Hall of Science. University of California, Berkeley.

Friedman, A. (Ed.). (2008). Framework for evaluating impacts of informal science education projects. Washington, DC: National Science Foundation.

Fusco, D. (2001). Creating relevant science through urban planning and gardening. Journal of Research in Science Teaching, 38(8), 860-877.

Gallagher, J., and Hogan, K. (2000). Intergenerational, community-based learning and science education. Journal of Research in Science Teaching, 37(2), 107-108.

Gee, J. (2003). What Video Games have to Teach Us about Learning and Literacy. New York: Palgrave Macmillan.

Gee, J. P. (2010). Human action and social groups as the natural home of assessment: Thoughts on 21st century learning and assessment. Innovative assessment for the 21st century: Supporting educational needs, 13-39.

Gee, J. P. (2010). New digital media and learning as an emerging area and“ worked examples” as one way forward. Mit Pr.

Gerber, B. L., Cavallo, A. M. ., & Marek, E. A. (2001). Relationships among informal learning environments, teaching procedures and scientific reasoning ability. International Journal of Science Education, 23(5), 535-549.

Gerber, B. L., Marek, E. A., & Cavallo, A. M. . (2001). Development of an informal learning opportunities assay. International Journal of Science Education, 23(6), 569-583.

Gibson, H., and Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students’ attitudes towards science. Science Education, 86(5), 693-705.

Goldman, S. (2006). “A New Angle on Families: Connecting the Mathematics in Daily Life with School Mathematics.” In Bekerman, Z., Burbules, N., Silberman-Keller, D. (Eds.), Learning in Places: The Informal Education Reader. Bern: Peter Lang.

Goldman, S. & Booker, A. (2009) Making Math a Definition of the Situation: Families as Sites for Mathematical Practices. Anthropology & Education Quarterly.

Goodwin, M.H. (2007). Occasioned knowledge exploration in family interaction. Discourse and Society, 18(1), 93-110.

Gration, M., and Jones, J. (2008, May/June). Learning from the process: Developmental evaluation within “agents of change.” ASTC Dimensions, From Intent to Impact: Building a Culture of Evaluation. Available: http://www.astc.org/ blog/2008/05/16/from-intent-to-impact-building-a-culture-of-evaluation/ [ac- cessed April 2009].

Gray, B. (2004). Informal learning in an online community of practice. Journal of Distance Education, 19(1), 20-35.

Green, J., & Luke, A. (2006). Introduction: rethinking learning: what counts as learning and what learning counts. Review of research in education, 30(1), xi-xiv.

Greenfield, P. M. (2009). Technology and Informal Education: What Is Taught, What Is Learned. Science, 323(5910), 69 -71.

Greenfield,P., & Lave,J. (1982). Cognitive aspects of informal education. In D. A. Wagner & H. W. Stevenson (Eds.), Cultural on child perspectives development (pp. 181-207). San Francisco: Freeman.

Gresalfi, M. S. (2009). Taking up Opportunities to Learn: Constructing Dispositions in Mathematics Classrooms. Journal of the Learning Sciences, 18(3), 43.

Griffin, J. (1994). Learning to learn in informal science settings. Research in Science Education, 24(1), 121-128.

Grossman, J. B., Goldmith, J., Sheldon, J., & Arbreton, A. J. A. (2009). Assessing after school settings. New Directions for Youth Development, 2009(121), 89-108.

Guichard, H. (1995). Designing tools to develop the conception of learners. International Journal of Science Education, 17(2), 243-253.

Gutnick, A. L., Robb, M., Takeuchi, L., Kotler, J. (2011). Always connected: The new digital media habits of young children.Joan Ganz Cooney Center at Sesame Workshop

Gutwill, J. P., & Allen, S. (2010). Facilitating family group inquiry at science museum exhibits. Science Education, 94(4), 710-742. doi:10.1002/sce.20387

Halpern, R., Spielberger, J., Robb, S., & Children, U. of C. C. H. C. for. (2001). Evaluation of the MOST (Making the Most of Out-of-School Time) Initiative: Final Report: Summary of Findings. Chapin Hall Center for Children, University of Chicago.

Halpern. R. (2009). The means to grow up: Re-inventing apprenticeship as a developmental support in adolescence. NY: Routledge.

Harris, E. , Deschenes, S., Wallace, A. (2011). Helping Older Youth Succeed Through Expanded Learning Opportunities. National Conference of State Legislatures & Harvard Family Research Project Brief Series, Brief no. 1.

Harris, E. (2012). Afterschool Evaluation 101: How to evaluate an expanded learning program

Harvard Family Research Project (2011). Helping Older Youth Succeed Through Expanded Learning Opportunities.

Harvard Family Research Project. “After School Programs in the 21st Century: Their Potential and What It Takes to Achieve It.” No. 10. Issues and Opportunities in Out-of-School Time Evaluation. Cambridge, MA: Harvard Family Research Project, February 2008.

Harvard Family Research Project. “Building and Evaluating Out-of-School Time Connections.” Vol. 12. No. 1 & 2. The Evaluation Exchange. Cambridge, MA: Harvard Family Research Project, Fall 2006.

Harvard Family Research Project. “Detangling Data Collection: Methods for Gathering Data.” No. 5. Out of School Time Evaluation Snapshot. Cambridge, MA: Harvard Family Research Project, August 2004.

Harvard Family Research Project. “Learning from Small-Scale Experimental Evaluations of After School Programs.” No. 8. Out-of-School Time Evaluation Snapshot. Cambridge, MA: Harvard Family Research Project, May 2006.

Harvard Family Research Project. “Measurement Tools for Evaluating Out-of-School Time Programs: An Evaluation Resource.” No. 6. Out-of-School Time Evaluation Snapshot. Cambridge, MA: Harvard Family Research Project, November 2005.

Harvard Family Research Project. “Performance Measures in Out-of-School Time Evaluation.” No.3. Out-of-School Time Evaluation Snapshot. Cambridge, MA: Harvard Family Research Project, March 2004.

Harvard Family Research Project. About the Harvard Family Research Project’s Out-of-School Time Program Evaluation Database. Cambridge, MA: Harvard Family Research Project.

Harvard Family Research Project. Moving Beyond Barriers – Attracting and Sustaining Youth Participation in Out-of-School Time Programs. No. 6. Cambridge, MA: Harvard Family Research Project, July 2004.

Hasan, R. (2002). Semiotic mediation and mental development in pluralistic societies: Some implications for tomorrow’s schooling. In G. Wells & G. Claxton (Eds.), Learning for life in the twenty-first cen tury: Sociocultural perspectives on the future of education (pp. 112–126). Oxford, UK: Blackwell.

Heath, S. (2000). Making learning work. Afterschool Matters, 1 (1), 33-45.

Heath, S., (2007). Diverse learning and learner diversity in “informal” science learning environments. Background paper for the Committee on Science Educa- tion for Learning Science in Informal Environments. Available: http://www7. nationalacademies.org/bose/Learning_Science_in_Informal_Environments_ Commissioned_Papers.html [accessed November 2008].

Herr-Stephenson, B., Rhoten, D., Perkel, D., & Sims, C. (2011). Digital Media and Technology in Afterschool Programs, Libraries, and Museums.

Hickey, D. T., Honeyford, M. A., Clinton, K. A., & McWilliams, J. (2010). Participatory Assessment of 21st Century Proficiencies. In V. Shute and B. J. Becker (Eds.) Innovative Assessment for the 21st Century: Supporting Educational Needs (pp.107-138). New York, New York: Springer.

Hirshfeld, L.A. (2002). Why anthropologists don't like children. American Anthropologist, 104(2), 611-627.

Hmelo-Silver, C. E., Katic, E., Nagarajan, A., & Chernobilsky, E. (2007). Soft leaders, hard artifacts, and the groups we rarely see: Using video to understand peer learning processes. In R. Goldman, R. D. Pea, B. Barron & S. J. Derry (Eds.), Video research in the learning sciences. Mahwah, NJ: Lawrence Erlbaum Associates.

Honey, M.A.& Hilton, M. (Eds). Committee on Science Learning: Computer Games, Simulations, and Education (2010). Learning science through computer games and simulations. Washington DC: The National Academies Press.

Honig, M., & McDonald, M. (2005). From promise to participation: Afterschool programs through the lens of socio-cultural learning theory. Afterschool Matters, Occasional Paper Series, (5).

Hsu, H. Y., & Wang, S. K. (2010). Using gaming literacies to cultivate new literacies. Simulation & Gaming, 41(3), 400-417.

Hull, G. A., & Katz, M. L. (2006). Crafting an agentive self: Case studies of digital storytelling. Research in the Teaching of English, 43-81.

Hull, G. A., & Nelson, M. E. (2005). Locating the semiotic power of multimodality. Written Communication, 22(2), 224.

Hull, G., & Schultz, K. (2001). Literacy and learning out of school: A review of theory and research. Review of educational research, 71(4), 575.

Hull, G.A., and Greeno, J.G. (2006). Identity and agency in nonschool and school worlds. In Z. Bekerman, N. Burbules, and D.S. Keller (Eds.), Learning in places: The informal education reader (pp. 77-97). New York: Peter Lang.

Isen, A.M. (2004). Some perspectives on positive feelings and emotions: Positive affect facilitates thinking and problem solving. In A.S.R. Manstead, N. Frijda, and A. Fischer (Eds.), Feelings and emotions: The Amsterdam symposium (pp. 263-281). New York: Cambridge University Press.

Ito, M., Baumer, S., Bittanti, M., boyd, danah, Cody, R., Herr-Stephenson, B., Horst, H. A., et al. (2009). Hanging Out, Messing Around, and Geeking Out: Kids Living and Learning with New Media (1st ed.). The MIT Press.

Ito, M., Horst, H., Bittanti, M., Boyd, D., Herr-Stephenson, B., Lange, P. G., Pascoe, C. J., et al. (2008). Living and learning with new media: Summary of findings from the Digital Youth Project. The John D. and Catherine T. MacArthur Foundation Reports on Digital Media and Learning.

Jackson, A., and Leahy, H.R. (2005). "Seeing it for real?" Authenticity, theater and learning in museums. Research in Drama Education, 10(3), 303-325.

Jarman, R. (2005). Science learning through scouting: An understudied context for informal science education. International Journal of Science Education, 27(4), 427-450.

Jenkins, H. (2009). Confronting the challenges of participatory culture: Media education for the 21st century. The MIT Press.

Jeong, A. (2005). A guide to analyzing message–response sequences and group interaction patterns in computer□mediated communication. Distance Education, 26(3), 367-383.

Jolly, E., Campbell, P., and Perlman, L. (2004). Engagement, capacity, continuity: A trilogy for student success. St. Paul: GE Foundation and Science Museum of Minnesota.

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. The Journal of the learning sciences, 4(1), 39-103.

Kafai, Y. (1995). Gender Differences in Children's Constructions of Video Games. In Interacting with Video. P. Greenfield and R. Cocking, eds. New York: Ablex Publishing. Pp. 39-52.

Kafai, Y. B. (2006). Playing and making games for learning: Instructionist and constructionist perspectives for game studies. Games and Culture 1(1): 34-40.

Kafai, Y. B. & Fields, D. A. (2009). Cheating in Virtual Worlds: Transgressive Designs for Learning. On the Horizon 17(1): 12-20.

Kafai, Y. B., Peppler, K. A., & Chiu, G. M. (2007). High Tech Programmers in Low-Income Communities: Creating a Computer Culture in a Community Technology Center. Communities and Technologies 2007, 545-563.

Kafai, Y.B. (1994). Minds in play: Computer game design as a context for children's learning. Mahwah, NJ: Lawrence Erlbaum Associates.

Kahne, J., Nagaoka, J., Brown, A., O'BRIEN, J., Quinn, T., & Thiede, K. (2001). Assessing after-school programs as contexts for youth development. Youth & Society, 32(4), 421.

Kane, T. J. (2004). The impact of after-school programs: Interpreting the results of four recent evaluations. New York: William T. Grant Foundation.

Kickmeier-Rust_digital_game_based_learning_bibliography

Kickmeier-Rust, M. D., & Albert, D. (2009). Emergence in digital educational games: A world of incidents in a universe of rules. In M. Pivec (Ed.), Proceedings of the 3rd European Conference on Games-based Learning (ECGBL), October 12-13, 2009, Graz, Austria.

Kickmeier-Rust, M. D., & Albert, D. (2009). Emergent design: Serendipity in digital educational games. In R. Shumaker (Ed.), Virtual and Mixed Reality, Lecture Notes in Computer Science (pp 206- 215). Berlin: Springer.

Kickmeier-Rust, M. D., & Albert, D. (2010). Micro Adaptivity: Protecting Immersion in Didactically Adaptive Digital Educational Games. Journal of Computer Assisted Learning, 26, 95-105.

Kickmeier-Rust, M. D., Hockemeyer, C., Albert, D., & Augustin, T. (2008). Micro adaptive, non-invassive assessment in educational games. In M. Eisenberg, Kinshuk, M. Chang, & R. McGreal (Eds.), Proceedings of the second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning (pp. 135-137), November 17-19, 2008, Banff, Canada.

Kickmeier-Rust, M. D., Marte, B., Linek, S. B., Lalonde, T., & Albert, D. (2008). Learning with computer games: Micro level feedback and interventions. In M. E. Auer (Ed.), Proceedings of the International Conference on Interactive Computer Aided Learning (ICL), September 24-27, 2008, Villach, Austria. Kassel: Kassel University Press.

Kickmeier-Rust, M. D., Steiner, C. M., & Albert, D. (2009). Non-invasive assessment and adaptive interventions in learning games. In Proceedings of the International Conference on Intelligent Networking and Collaborative Systems (INCoS), November 4-6, 2009, Barcelona, Spain.

Kirriemuir, J., & McFarlane, A. (2004). Literature review in games and learning.

Klopfer, E. (2008). Augmented Learning; Research and Design of Mobile Educational Games. Massachusetts: The MIT Press.

Koran, J.J., Koran, M.L., and Foster, J.S. (1988). Using modeling to direct attention. Curator, 31(1), 36-42.

Kow, Y. M. and Nardi, B. 2011. Forget Online Communities? Revisit Cooperative Work! Proceedings Conference on Computer-supported Cooperative Work.

Kow, Y.M. and Nardi, B. (2009). Culture and Creativity: World of Warcraft Modding in China and the U.S. In B. Bainbridge, ed. Online Worlds: Convergence of the Real and the Virtual. Springer.

Lave, J. (1996). Teaching, as learning, in practice. Mind, Culture, and Activity, 3(3), 149-164.

Law, E., Kickmeier-Rust, M., Albert, D., & Holzinger, A. (2008). Challenges in the development and evaluation of immersive digital educational games. HCI and Usability for Education and Work, 19-30.

Learning Point Associates. (2006). 21st century community learning centers (21st CCLC) analytic support for evaluation and program monitoring: An overview of the 21st CCLC program: 2004-05. Naperville, IL: Author.

Learning Science in Informal Environments Resources

Lee, C. (2003). Toward a framework for culturally responsive design in multimedia computer environments: Cultural modeling as a case. Mind, Culture, and Activ- ity, 10(2), 42-61.

Lee, O., and Fradd, S. (1996). Interactional patterns of linguistically diverse students and teachers: Insights for promoting science learning. Linguistics and Educa- tion, 8(3), 269-297.

Leinhardt, G., & Knutson, K. (2004). Listening in on museum conversations. Rowman Altamira.

Leinhardt, G., Crowley, K., & Knutson, K. (2002). Learning conversations in museums. Psychology Press.

Leinhardt, G., Tittle, C., & Knutson, K. (2002). Talking to oneself: Diaries of museum visits. Learning conversations in museums, 103-133.

Leinhardt, G., Tittle, C., and Knutson, K. (2000). Talking to Oneself: Diaries of Museum Visits.

Lemke, J. (2011) Appropriate Assessment in a Living Activity System. LCHC internal document

Little, P. and Lauver, S. “Engaging Adolescents in Out-of-School Time Programs: Learning What Works.” Vol. 12. No. 2. The Prevention Researcher. Eugene, OR: Integrated Research Services, Inc., 2005.

Little, P. and Lauver, S. “Finding the Right Hook: Strategies for Attracting and Sustaining Participation in After-School Programs.” School Administrator. Arlington, VA: American Association of School Administrators, May 2005.

Little, P., & Harris, E. (2003). A Review of Out-of-School Time Program Quasi-Experimental and Experimental Evaluation Results. Out-of-School Time Evaluation Snapshot.

Little, P., Dupree, S. and Deich, S. (2002). Documenting Progress and Demonstrating Results: Evaluating Local Out-of-School Time Programs. New York, NY: The Finance Project.

Livingstone, D. W. (2000). Exploring the icebergs of adult learning: Findings of the first Canadian survey of informal learning practices.

Loh, C. S., Anantachai, A., Byun, J., & Lenox, J. (2007). Assessing what players learned in serious games: in situ data collection, information trails, and quantitative analysis. 10th International Conference on Computer Games: AI, Animation, Mobile, Educational & Serious Games (CGAMES 2007), July (Vol. 2528).

Lu, S., and Graesser, A.C. (in press). An eye tracking study on the roles of texts, pictures, labels, and arrows during the comprehension of illustrated texts on device mechanisms. Submitted to Cognitive Science.

Macdonald, J. (2003). Assessing online collaborative learning: process and product. Computers & Education, 40(4), 377-391.

Mahiri, J. (2011). Digital Tools in Urban Schools: Mediating a Remix of Learning. University of Michigan Press.

Mahoney, J., Larson, R., Eccles, J. S. (Eds.) (2005). Organized activities as contexts of development: Extracurricular activities, after-school, and community programs. Hillsdale, NJ: Erlbaum.

Maloney, J. H., Peppler, K., Kafai, Y., Resnick, M., & Rusk, N. (2008). Programming by choice: urban youth learning programming with scratch. ACM SIGCSE Bulletin, 40(1), 367-371.

Martin, L., & Goldman, S. (2010). Family inheritance: parallel practices of financial responsibility in families. In L. Lin, H. Varenne, & E.W. Gordon (Eds.), Educating comprehensively: varieties of educational experiences (Vol. 3). Lewiston, NY: The Edwin Mellen Press.

Martin, L., and Toon, R. (2005). Narratives in a science center: Interpretation and identity. Curator, 48(4), 407-426.

Martin, L., Goldman, S., & Jimenez, O. (2009) The Tanda: A Practice at the Intersection of Mathematics, Culture, and Financial Goals." Mind, Culture & Activity, 16(4).

Martin, L.M. (2004). An emerging research framework for studying informal learning and schools. Science Education, 88 (Suppl. 1), S71-S82.

Maxwell, L.E., and Evans, G.W. (2002). Museums as learning settings: The importance of the physical environment. Journal of Museum Education, 27(1), 3-7.

Mayer, R. E., Schustack, M. W., & Blanton, W. E. (1999). What do children learn from using computers in an informal, collaborative setting? EDUCATIONAL TECHNOLOGY-SADDLE BROOK NJ-, 39, 27-31.

Mayer, R.E. (1997). Multimedia learning: Are we asking the right questions? Educational Psychologist, 32(1), 1-19.

Mayer, R.E., Quilici, J., Moreno, R., Duran, R., Woodbridge, S., Simon, R., Sanchez, D., and Lavezzo, A. (1997). Cognitive consequences of participation in a fifth dimension after-school computer club. Journal of Educational Computing Re- search, 16(4), 353-369.

McLaughlin, M. W. (2000). Community Counts: How Youth Organizations Matter for Youth Development.

McPherson, T. (2008). Digital youth, innovation, and the unexpected. The MIT Press.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies.

Meyer, R. L., Bevan, B., & Garza, P. (n.d.). Museums Afterschool: Principles, Data, and Design.

Michalchik, V., & Gallagher, L. (2010). Naturalizing Assessment. Curator: The Museum Journal, 53(2), 209-219.

Miller, B. M. (2001). The promise of after-school programs. Educational Leadership, 58, n7.

Miller, B. M. (2003). Critical Hours: Afterschool Programs and Educational Success. Quincy, MA: Nellie Mae Education Foundation.

Miller, S. (1971). Ends, Means and Galumphing: Some Leitmotifs of Play. American Anthropologist 75: 87-98.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). A brief introduction to evidence- centered design (CSE Report 632). CA: Center for Research on Evaluation, Standards, and Student Testing. (ERIC Document Reproduction Service No. ED483399)

Mitchell, A. and Savill-Smith C. (2005). The use of computer and video games for learning. A review of the literature. London. Learning and Skills Development Agency.

Moje, E. B., & Tysvaer, N. (2010). Adolescent literacy development in out-of-school time: A practitioner's guide. New York, NY: Carnegie Corporation of New York.

Moje, E. B., Eccles, J. S. (2005). Out-of-school programs for literacy development review of the literature. New York Adolescent Literacy Council, Carnegie Corporation.

Moll, L., Amanti, C., Neff, D., and Gonzalez, N. (2005). Funding of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. In L. Moll, C. Amanti, and N. Gonzalez (Eds.), Funds of knowledge: Theorizing practices in households, communities, and classrooms (pp. 71-88). London: Routledge.

Moussouri, T. (1997). The use of children's drawings as an evaluation tool in the museum. Museological Review, 4, 40-50.

Naismith, L., Lonsdale, P., Vavoula, G., & Sharples, M. (2004). Literature Review in Mobile Technologies and Learning: Report 11. Educational Technology.

Nardi, B., Ly, S. and Harris, J. (2007). Learning conversations in World of Warcraft. Proceedings Hawaii International Conference on Systems Science.

Nasir, N. S. (2005). Individual cognitive structuring and the sociocultural context: Strategy shifts in the game of dominoes. The Journal of the Learning Sciences, 14(1), 5-34.

Nasir, N.S., Rosebery, A.S., Warren B., and Lee, C.D. (2006). Learning as a cultural process: Achieving equity through diversity. In R.K. Sawyer (Ed.), The Cam- bridge handbook of the learning sciences (pp. 489-504). New York: Cambridge University Press.

National Institute on Out-of-School Time and the Academy for Educational Development Center for Youth Development and Policy Research. Strategic Plan: Building a Skilled and Stable Out-of-School Time Workforce. Wellesley, MA: NIOST, September 2003.

National Research Council. (2001). Knowing what students know: The science and design of educational assessment. Committee on the Foundations of Assessment. J.W. Pellegrino, N. Chudowsky, and R. Glaser (Eds.). Washington, DC: National Academy Press.

National Research Council. (2009). Learning science in informal environments: People, places, and pursuits. Washington, DC: The National Academy Press.

Nelson-Barber, S., and Estrin, E. (1995). Culturally responsive mathematics and sci- ence education for Native students. Washington, DC: Native Education Initiative of the Regional Educational Labs.

Neulight, N., Kafai, Y.B., Kao, L., Foley, B., and Galas, C. (2007). Children's participation in a virtual epidemic in the science classroom: Making connections to natural infectious diseases. Journal of Science Education and Technology, 16(1), 47-58.

New York State Afterschool Network. Program Quality Self-Assessment Tool. Planning for Ongoing Program Improvement. Albany, NY: NYSAN, January 6, 2005.

Newman, Griffin Cole (1989) The Construction Zone

Nixon, A. S. (2009). Mediating social thought through digital storytelling. Pedagogies: An International Journal, 4(1), 63-76.

Noam, G. G. (2004). (Ed.). After-school worlds: Creating a new social space for development and learning. New Directions for Youth Development, Number 101.

Nocon, H. D. (2004). Sustainability as process: community education and expansive collaborative activity. Educational Policy, 18(5), 710.

O'Connor, K., & Penuel, W. R. (2010). Introduction: Principles of a human sciences approach to research on learning. The Yearbook of the National Society for the Study of Education, 109, 1-16.

O'Neill, M.C., and Dufresne-Tasse, C. (1997). Looking in everyday life/Gazing in museums. Museum Management and Curatorship, 16 (2), 131-142.

Oblinger, D. G. (2006). Learning Spaces | EDUCAUSE. Educase. Retrieved from http://www.educause.edu/LearningSpaces

Osborne, J., & Dillon, J. (2007). Research on learning in informal contexts: Advancing the field?

Out-Of-School Time Programs In Rural Areas

Partee, G. Lessons Learned About Effective Policies and Practices for Out-of-School Time Programming. Washington, DC: American Youth Policy Forum, 2003.

Partnership for 21st Century Skills (2004). 21st Century Readiness for Every Child. Tucson, AZ.

Peirce N., Conlan O., Wade V. (2008). Adaptive Educational Games: Providing Non-invasive Personalised Learning Experiences. In: Digital Games and Intelligent Toys Based Education, 2008 Second IEEE International Conference on (2008), pp. 28-35.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). Knowing what students know: The science and design of educational assessment. National Academies Press.

Penuel, W Defining a Model for Collaborative Research at the Intersection of Learning, Reform Implementation, and Design

Peppler, K. A., & Kafai, Y. B. (2007). From SuperGoo to Scratch: exploring creative digital media production in informal learning. Learning, Media and Technology, 32(2), 149-166.

Perry, D. L. (1993). Measuring learning with the knowledge hierarchy. Visitor Studies: Theory, Research and Practice: Collected Papers from the 1993 Visitor Studies Conference, 6, 73-77.

Phipps, M. (2010). Research Trends and Findings From a Decade (1997–2007) of Research on Informal Science Education and Free-Choice Science Learning. Visitor Studies, 13(1), 3-22.

Pinkard, Nichole. YouMedia & iRemix

Pivec M (Ed.), (2006) Affective and emotional aspects of human-computer interaction; Game-Based and Innovative Learning Approaches. Vol.1: The Future of Learning (IOS Press, 2006).

Plass, J. L., Homer, B. D., & Hayward, E. O. (2009). Design factors for educationally effective animations and simulations. Journal of Computing in Higher Education, 21(1), 31-61.

Policy Studies Associates. “Identifying Staffing Needs and Recruiting Qualified After- School Staff.” The After-School Corporation Research Brief. Washington, DC: PSA, October 2000.

Policy Studies Associates. “Matching After-School Staff to Project Needs and Activities.” The After-School Corporation Research Brief. Washington, DC: PSA, October 2000.

Polman, J. L. (2006). Mastery and appropriation as means to understand the interplay of history learning and identity trajectories. The Journal of the Learning Sciences, 15(2), 221-259.

Posner, J. K., & Vandell, D. L. (1994). Low Income Children’s After School Care: Are There Beneficial Effects of After School Programs? Child development, 65(2), 440-456.

Prensky, M. (2000). Digital game-based learning. New York: McGraw-Hill.

Rahm, J. (2002). Emergent learning opportunities in an inner□city youth gardening program. *Journal of Research in Science Teaching, 39*(2), 164-184.

Randol, S.M. (2005). The nature of inquiry in science centers: Describing and assessing inquiry at exhibits. Unpublished doctoral dissertation, University of California, Berkeley.

Reisner, E. R., White, R. N., Russell, C. A., & Birmingham, J. (2004). Building quality, scale and effectiveness in after-school programs. Report submitted to The After-School Corporation. Washington, DC: Policy Studies Associates, Inc.

Rennie, L. J., & Johnston, D. J. (2004). The nature of learning and its implications for research on learning from museums. Science Education, 88(S1), S4-S16.

Rennie, L. J., Feher, E., Dierking, L. D., & Falk, J. H. (2003). Toward an agenda for advancing research on science learning in out□of□school settings. Journal of Research in Science Teaching, 40(2), 112-120.

Rennie, L.J., and McClafferty, T.P. (1993). Learning in science centres and science museums: A review of recent studies. Research in Science Education, 23(1), 351.

Renninger, K., Hidi, S., and Krapp, A. (1992). The role of interest in learning and development. Mahwah, NJ: Lawrence Erlbaum Associates.

Resnick, L. B. (1987). The 1987 presidential address: Learning in school and out. Educational researcher, 16(9), 13-54.

Resources On Afterschool 2009 Bibliography

Rhodes, J. E. (2004). The critical ingredient: Caring youth□staff relationships in after□school settings. New Directions for Youth Development, 2004(101), 145-161.

Robinson, G. and Fenwick, L. More than homework, a snack, and basketball: Afterschool Programs as an Oasis of Hope for Black Parents in Four Cities. Washington, DC: Black Alliance Educational Options, October 2007.

Rockman Et Al. (2007). Media-based learning science in informal environments. Background paper for the Learning Science in Informal Environments Committee of the National Research Council. Available: http://www7.nationalacademies.org/ bose/Rockman_et%20al_Commissioned_Paper.pdf [accessed October 2008].

Rogoff, B., Paradise, R., Mejía Arauz, R., Correa-Chávez, M., and Angelillo, C. (2003). Firsthand learning by intent participation. Annual Review of Psychology, 54, 175-203.

Ross, N., Medin, D., Coley, J.D., and Atran, S. (2003). Cultural and experiential differences in the development of folkbiological induction. Cognitive Development, 18(1), 35-47.

Roth, J. L., Malone, L. M., & Brooks-Gunn, J. (2010). Does the amount of participation in afterschool programs relate to developmental outcomes? A review of the literature. American journal of community psychology, 45(3), 310-324.

Rounds, J. (1999). Meaning-making: A new paradigm for museum exhibits? Exhibitionist, 18(2), 5-8.

Roussou, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C., and Barnes, C. (1999). Learning and building together in an immersive virtual world. Presence: Teleop- erators and Virtual Environments Journal, 8(3), 247-263.

Roussou, M., Oliver, M., & Slater, M. (2006). The Virtual Playground: an Educational Virtual Reality Environment for Evaluating Interactivity and Conceptual Learning. Journal of Virtual Reality 10(3-4), Springer, pp. 227-240.

Roussou, M., Oliver, M., & Slater, M. (2007) Exploring Activity Theory as a Tool for Evaluating Interactivity and Learning in Virtual Environments for Children. Journal of Cognition, Technology & Work, Springer, ISSN: 1435-5558 (Print) 1435-5566 (Online).

Sabo-Flores, K. (2003). A Vygotskian perspective on youth participatory evaluation. Youth participatory evaluation: A field in the making, Special Issue of New Directions for Evaluation, no. 98, 13-24.

Sabo-Flores, K. (2009). A dynamic framework for understanding the complex work of quality out-of-school programs. New York: The Robert Bowne Foundation.

Salen, K. (2008). The ecology of games: Connecting youth, games, and learning. The MIT Press.

Salen, K., Torres, R., Wolozin, L., Rufo-Tepper, R., & Shapiro, A. (2010). Quest to Learn: Developing the School for Digital Kids. The MIT Press.

Sandoval, W.A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. Science Education, 89(4), 634-656.

Scaffidi, C., & Chambers, C. (2011). Skill progression demonstrated by users in the Scratch animation environment. International Journal of Human-Computer Interaction

Scanlon, E., Jones, A., and Waycott, J. (2005). Mobile technologies: Prospects for their use in learning in informal science settings. Journal of Interactive Media in Education, 21(5), 1-17.

Schauble, L., Leinhardt, G., and Martin, L. (1998). Organizing a cumulative research agenda in informal learning contexts. Journal of Museum Education, 22(2 and 3), 3-7.

Schmidt, R. (2010). Building a Badge Framework to Ignite Student Learning Identities. Digital Youth Network. Bill and Melinda Gates Foundation in collaboration with the MacArthur Foundation

Schwartz, D. L., & Noam, G. Informal Science Learning in Afterschool Settings: A Natural Fit

Schwartz, D. L., & Arena, D. (2009). Choice-based assessments for the digital age. MacArthur 21st Century Learning and Assessment Project.

Schwier, R.A. (2010). Focusing educational technology research on informal learning environments. Contemporary Educational Technology, 1(1), 90-92.

Scott-Little, C., Hamann, M. S., & Jurs, S. G. (2002). Evaluations of after-school programs: A meta-evaluation of methodologies and narrative synthesis of findings. American Journal of Evaluation, 23(4), 387-419.

Scribner, S., & Cole, M. (1973). Cognitive consequences of formal and informal education. Science, 182(4112), 553-559.

Seal-Wanner, C. (2007). eTEENS: Teens and Technology: The perfect storm? Television Quarterly, 37(2), 5-16.

Sefton-Green, J. (2003). Informal Learning: substance or style? Teaching Education, 14(1), 37-51.

Sefton-Green, J. (2004). Literature review in informal learning with technology outside school.

Sefton-Green, J. (2006). Youth, technology, and media cultures. Review of research in education, 30, 279-306.

Shaffer, D. W. & Gee, J. P. EAGER Proposal for Research in Measurement and Modeling: Dynamic STEM Assessment through Epistemic Network Analysis.

Shaffer, D. W., & Gee, J. P. (2012). The Right Kind of GATE: Computer games and the future of assessment.. In M. Mayrath, J. Clarke-Midura, & D. H. Robinson (Eds.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research*. Charlotte, NC: Information Age Publishing.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., et al. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning. International Journal of Learning, 1(2), 33-53.

Shaffer, D.W. (2006). How computer games help children learn. New York: Palgrave Macmillan.

Shepard, L. (2000). The role of assessment in a learning culture. Educational Researcher, 29(7), 4-14.

Shernoff, D. J., & Vandell, D. L. (2007). Engagement in after-school program activities: Quality of experience from the perspective of participants. Journal of Youth and Adolescence, 36(7), 891-903.

Shute, V. J. (2009). Simply assessment. International Journal of Learning and Media, 1(2), 1-11.

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. Computer games and instruction. Charlotte, NC: Information Age Publishers.

Shute, V. J. & Becker, B. J. (2010). Innovative assessment for the 21st century: Supporting educational needs (pp. 1-11). New York, NY: Springer-Verlag.

Shute, V. J. & Becker, B. J. (2010). Prelude: Issues and assessment for the 21st century. In V. J. Shute & B. J. Becker (Eds.), Innovative assessment for the 21st century: Supporting educational needs (pp. 1-11). New York, NY: Springer-Verlag.

Shute, V. J., & Torres, R. (2012). Where streams converge: Using evidence-centered design to assess Quest to Learn. Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research. Charlotte, NC: Information Age Publishing.

Shute, V. J., Masduki, I., Donmez, O., Dennen, V. P., Kim, Y. J., Jeong, A. C., & Wang, C. Y. (2010). Modeling, assessing, and supporting key competencies within game environments. Computer-based diagnostics and systematic analysis of knowledge, 281-309.

Shute, V. J., Ventura, M., Bauer, M. I., & Zapata-Rivera, D. (2009). Melding the pow- er of serious games and embedded assessment to monitor and foster learn- ing: Flow and grow. In U. Ritterfeld, M. Cody, & P. Vorderer (Eds.), Serious games: Mechanisms and effects (pp. 295–321). Mahwah, NJ: Routledge, Taylor and Francis.

Simkins, D. & Steinkuehler, C. (2008). Critical ethical reasoning & role play. Games & Culture, 3, (333-355).

Simon, H.A. (2001). “Seek and ye shall find”: How curiosity engenders discovery. In K. Crowley, C. Schunn, and T. Okada (Eds.), Designing for science: Implica- tions from everyday, classroom, and professional settings (pp. 5-20). Mahwah, NJ: Lawrence Erlbaum Associates.

Siyahhan, S., Barab, S., & Downton, M. (2010). Using activity theory to understand intergenerational play: The case of Family Quest. International Journal of Computer-supported Collaborative Learning, 5(4), 415-432.

Smith, B.K., Sharma, P., & Hooper, P. (2006). Decision making in online fantasy sports communities. Interactive Technology and Smart Education, 3(4), 347 - 360

Spielberger, J., & Halpern, R. (2002). The role of after-school programs in children’s literacy development. Chicago: University of Chicago, Chapin Hall Center for Children.

Squire, K. (2003). Video games in education. International Journal of Intelligent Games and Simulation, 2(1), 49-62.

Squire, K. (2005). Changing the game: What happens when video games enter the classroom. Innovate: Journal of online education, 1(6).

Squire, K. (2006). From content to context: Videogames as designed experience. Educational Researcher, 35(8), 19-29.

Squire, K. (2008). Open-ended video games: A model for developing learning for the interactive age. In K. Salen (Ed.) The John D. and Catherine T. MacArthur Foundation series on digital media and learning. (167-198) Cambridge, MA: The MIT Press.

Squire, K., and Patterson, N. (2009). Games and simulations in informal science education. Paper commissioned for the National Research Council Workshop on Gaming and Simulations, October 6-7, Washington, DC.

Squire, K., and Patterson, N. (2009). Games and simulations in informal science education. Paper commissioned for the National Research Council Workshop on Gaming and Simulations, October 6-7, Washington, DC. Available: http://www7. nationalacademies.org/bose/Gaming_Sims_Commissioned_Papers.html [accessed March 12, 2010].

Squire, K.D. (2005). Resuscitating research in educational technology: Using game-based learning research as a lens for looking at design-based research. Educational Technology 45(1), 8-14.

Squire, K.D. (2005). Toward a theory of games literacy. Telemedium 52 (1-2), 9-15.

Squire, K.D., Jan, M., Matthews, J., Wagler, M., Martin, J., Devane, B. & Holden, C. (2007). Wherever you go, there you are: The design of local games for learning. In B. Sheldon & D. Wiley (Eds). The design and use of simulation computer games in education, (265-296). Rotterdam, Netherlands: Sense Publishing.

Stake, J.E., and Mares, K.R. (2005). Evaluating the impact of science-enrichment programs on adolescents' science motivation and confidence: The splashdown effect. Journal of Research in Science Teaching, 42(4), 359-375.

Steinkuehler, C. (2004). Learning in massively multiplayer online games. Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.). Proceedings of the Sixth International Conference of the Learning Sciences, (521–528). Mahwah, NJ: Erlbaum.

Steinkuehler, C. (2008). Cognition and literacy in massively multiplayer online games. In D. Leu, J. Coiro, C. Lankshear, & K. Knobel (Eds.). Handbook of Research on New Literacies, (pp. 611- 634). Mahwah NJ: Erlbaum.

Steinkuehler, C. (2008). Massively multiplayer online games as an educational technology: An outline for research. Educational Technology, 48(1), (10-21).

Steinkuehler, C. & Duncan, S. (in press). Informal scientific reasoning in online virtual worlds. Journal of Science Education & Technology.

Stevens, R. (2007). Capturing ideas in digital things: the traces digital annotation medium. In R. Goldman, B. Barron, and R. Pea (Eds.), Video research in the learning sciences. Cambridge: Cambridge University Press.

Stevens, R. (2010). Learning as a members' phenomenon: Toward an ethnographically adequate science of learning. NSSE Yearbook, 109(1).

Stevens, R., and Toro-Martell, S. (2003). Leaving a trace: Supporting museum visitor interpretation and interaction with digital media annotation systems. Journal of Museum Education, 28(2), 25-31.

Stevens, R., Satwicz, T., & McCarthy, L. (2008). In game, In room, In world: Reconnecting video game play to the rest of kids' lives. The Ecology of Games. Salen, K. (Ed.), Cambridge: MIT Press.

Stevens, Reed. YouSTEAM

Stocklmayer, S. M., Rennie, L. J., & Gilbert, J. K. (2010). The roles of the formal and informal sectors in the provision of effective science education. Studies in Science Education, 46(1), 1-44.

Stone, L. D., & Gutiérrez, K. D. (2007). Problem articulation and the processes of assistance: An activity theoretic view of mediation in game play. International Journal of Educational Research, 46(1-2), 43-56.

Straka, G. A. (2004). Informal learning: genealogy, concepts, antagonisms and questions. Bremen: Univeristät Bremen, ITB Forschungsberichte, 15, 2004.

Strobel, K. Kirshner, B., O'Donoghue, J. and McLaughlin, M. (2008). Qualities that attract urban youth to after-school settings and promote continued participation. Teachers College Record, 110-8, 1177-1705.

Tenenbaum, H.R., and Callanan, M.A. (2008). Parents' science talk to their children in Mexican-descent families residing in the United States. International Journal of Behavioral Development, 32(1), 1-12.

Tharp, R.G., and Gallimore, R. (1989). Rousing minds to life: Teaching and learning in social context. New York: Cambridge University Press.

Thompson, C.C. and Diaz, L.B. (2009). Building Identities as Experts: Youth Learning in an Urban After-School Space, In Cynthia Carter Ching and Brian Foley (Eds.), Technology and Identity: Constructing the Self in a Digital World. New York, NY: Cambridge University Press.

Thompson, S., and Bonney, R. (2007, March). Evaluating the impact of participation in an on-line citizen science project: A mixed-methods approach. In J. Trant and D. Bearman (Eds.), Museums and the web 2007: Proceedings. Toronto: Archives and Museum Informatics. Available: http://www.archimuse.com/mw2007/papers/ thompson/thompson.html [accessed October 2008].

Tseng, V., & Seidman, E. (2007). A systems framework for understanding social settings. American Journal of Community Psychology, 39(3), 217-228.

Vadeboncoeur, J. (2006). Engaging young people: Learning in informal contexts. Review of Research in Education, 30, 239-278.

Vandell, D. and Reisner, E. Study of Promising Afterschool Programs: Key Findings from New Research on the Benefits of High-Quality Afterschool Programs. University of California-Irvine, University of Wisconsin and Policy Studies Associates, October 2007.

Vandell, D. L., Reisner, E. R., & Pierce, K. M. (2007). Outcomes linked to high-quality afterschool programs: Longitudinal findings from the study of promising afterschool programs. University of California, Irvine and Policy Studies Associates, Inc.

Vandell, D. L., Shernoff, D. J., Pierce, K. M., Bolt, D. M., Dadisman, K., & Brown, B. B. (2005). Activities, engagement, and emotion in after school programs (and elsewhere). New Directions for Youth Development, 2005(105), 121-129.

Vandenberg, B. (1998). Real and Not Real: A Vital Developmental Dichotomy. In Multiple Perspectives on Play in Early Childhood Education. O. Saracho and B. Spoedek, eds. Albany: SUNY Press.

Vasquez, O. A. (2003). La clase magica. Lawrence Erlbaum.

vom Lehn, D., Heath, C., and Hindmarsh, J. (2002). Video-based field studies in museums and galleries. Visitor Studies, 5 (3), 15-23.

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. Journal of Educational Psychology, 102, 860-871.

Warren, B., Ballenger, C., Ogonowski, M., Rosebery, A., and Hudicourt-Barnes, J. (2001). Rethinking diversity in learning science: The logic of everyday sense- making. Journal of Research in Science Teaching, 38, 529-552.

Watson, J. (1998). If you don't have it, you can't find it: A close look at students' perceptions of using technology. Journal of the American Society for Information Science, 49(11), 1024-1036.

Weiss, H., & Little, P. (2003). Why, when, and how to use evaluation: experts speak out. Issues and Opportunities in Out-of-School Time Evaluation. Cambridge, MA: Harvard Family Research Project.

Whitehead J., McLemore, B. and Orlando, M. (2008). World of Warcraft Programming. New York: Wiley

Wiley Periodicals, Inc. "Participation in Youth Programs: Enrollment, Attendance, and Engagement." New Directions for Youth Development. Issue 105. Hoboken, NJ: Wiley Periodicals, Inc, Fall 2005.

William T. Grant Foundation (2011). From Soft Skills to Hard Data: Measuring Youth Program Outcomes.

Wimer, C. and Hull, B. Harnessing Technology in Out-of-School Time Settings. Cambridge, MA: Harvard Family Research Project, January 2006.

Yohalem, N. (2003). Adults who make a difference. Community youth development: Programs, policies, and practices, 358-372.

Youth Program Quality Assessment (instrument)

YouthLearn Initiative. Model Technology Integration in Afterschool. Newton, MA: Education Development Center.

Zief, S. G. (2004). Impacts of after-school programs on student outcomes. University of Pennsylvania.

Appendix C: Online Resources: Assessment, Funding, Research

Afterschool Alliance http://www.afterschoolalliance.org/
Afterschool Resources.org http://afterschoolresources.org/
Annie E. Casey Foundation - http://www.aecf.org/
Assessment Tools in Informal Science (ATIS). http://www.pearweb.org/atis/
Bill & Melinda Gates Foundation http://www.gatesfoundation.org/Pages/home.aspx
California Academy of Sciences http://www.calacademy.org/
Center For Assessment. http://www.nciea.org/
Center for Informal Learning and Schools http://cils.exploratorium.edu/cils/page.php?ID=23
Coalition for Science After School http://www.afterschoolscience.org/
Digital Media Learning Central http://dmlcentral.net/
Encyclopedia of informal education http://www.infed.org/index.htm
Exploratorium http://www.exploratorium.edu/
Futurelab http://www.futurelab.org.uk/
Harvard Family Research Project http://www.hfrp.org/
Humanities, Arts, Science, and Technology Advanced Collaboratory http://hastac.org/
Informal Learning Review http://www.informallearning.com/ilr-about.htm
Informal Science Education Evidence Wiki http://iseevidencewiki.org/index.php/Main_Page
Informal Science : informal learning projects, research and evaluation. http://informalscience.org/
Institute of Play http://www.instituteofplay.org/
Institute for Learning Innovation http://www.ilinet.org/display/ILI/Home
iRemix http://iremix.org/
John D. & Catherine T. MacArthur Foundation http://www.macfound.org/
Journal of Science Education and Technology http://www.springerlink.com/content/v8h0x482136u648m/
National-Academies.org http://www.nationalacademies.org/
National Center for the Improvement of Educational Assessment http://www.nciea.org/
National Research Council http://sites.nationalacademies.org/NRC/index.htm
National Science Foundation Informal Science Education
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5361
National Science Foundation Innovative Technology Experiences for Students and Teachers (ITEST) program online resource library http://itestlrc.edc.org/resource_library
New Media Literacies Community Site http://projectnml.ning.com/

Partnership for 21st Century Skills - http://www.p21.org/
PEAR - Program in Education, Afterschool & Resiliency: Informal Science Education assessment database. http://www.pearweb.org/
Relating Research to Practice: Current Research in Science Education http://www.research2practice.info/
Remediating Assessment blog http://remediatingassessment.blogspot.com
Transformative Learning Technologies Laboratory. http://tltl.stanford.edu/fablearn2011
U.S. Department of Education http://www.ed.gov/
U.S. Department of Education, Office of Educational Technology
http://www2.ed.gov/about/offices/list/os/technology/index.html
U.S. Department of Education - Reports http://www2.ed.gov/about/pubs/publications-reports.html
William T Grant Foundation http://www.wtgrantfoundation.org/Home

Examples Out-of-School Learning Projects

Computer Clubhouse Network. http://www.computerclubhouse.org/
Digital Youth Network: http://digitalyouthnetwork.org/
Elektra Project http://www.elektra-project.org/
FabLab@School http://tltl.stanford.edu/projects/fablabschool
Games For Learning Institute. http://g4li.org/
Healthy Youth Peer Education Program (H.Y.P.E.) http://www.hypeprogram.org/HYPE/H.Y.P.E..html
JILA PFC PISEC | CU-Boulder. http://www.colorado.edu/physics/PISEC/
LIFE Center: Learning in Informal and Formal Environments http://life-slc.org/
Quest to Learn http://www.q2l.org/
Stanford Makers' Club http://stanfordmakersclub.ning.com/
University-Community Links. http://uclinks.berkeley.edu/
You Media. http://youmediachicago.org/
YouSTEM: Interest-Driven STEM Exploration Space. http://osep.northwestern.edu/projects/youstem
Youth Media International, Youth Radio. http://www.youthradio.org/