Epistemic Network Analysis: A Prototype for 21st Century Assessment of Learning

IJLM

[AQ1] **David Williamson Shaffer**
University of Wisconsin-Madison

David Hatfield
University of Wisconsin-Madison

Gina Navoa Svarovsky
University of Wisconsin-Madison

Padraig Nash
University of Wisconsin-Madison

Aran Nulty
University of Wisconsin-Madison

Elizabeth Bagley
University of Wisconsin-Madison

Ken Frank
Michigan State University

André A. Rupp
University of Maryland

Robert Mislevy
University of Maryland

[AQ2] **Keywords**

Visit IJLM.net

doi: 10.1162/ijlm.2009.0013

Volume 1, Number 2

Abstract

In this article we examine educational assessment in the 21st century. Digital learning environments emphasize learning in action. In such environments, assessments need to focus on performance in context rather than on tests of abstracted and isolated skills and knowledge. Digital learning environments also provide the potential to assess performance in context, because digital tools make it possible to record rich streams of data about learning in progress. But what assessment methods will use this data to measure mastery of complex problem solving—the kind of thinking in action that takes place in digital learning environments?

Here we argue that one way to address this challenge is through *evidence-centered design*[1]—a framework for developing assessments by systematically linking models of understanding, observable actions, and evaluation rubrics to provide evidence of learning. We examine how evidence-centered design can address the challenge of assessment in new media learning environments by presenting one specific theory-based approach to digital learning, known as *epistemic games* (http://epistemicgames.org/eg/), and describing a method, *epistemic network analysis (ENA)*, to assess learner performance based on this theory. We use the theory and its related assessment method to illustrate the concept of a *digital learning system*—a system composed of a theory of learning and its accompanying method of assessment, linked into an evidence-based, digital intervention. We argue that whatever tools of learning and assessment digital environments use, they need to be integrated into a coherent digital learning system linking learning and assessment through evidence-centered design.

The Challenges of a New Century

In this article we look at educational assessment in the 21st century. Digital media create new models, approaches, and techniques for learning—as well as new educational outcomes, goals, and needs (Papert 1980; Shaffer 2007; Shaffer, Squire, Halverson, and Gee, 2005a) In an age of global competition brought about in part by digital media, economic success, cultural literacy, and civic participation require more than the traditional basic facts and basic skills emphasized in current curricula and standardized tests.[2] Digital literacy requires new kinds of thinking and digital media provides new ways of learning. Thus digital education creates new challenges for assessment (Gee 2003; Gee, Hull, and Lankshear 1996; Shaffer 2007).

Digital learning environments emphasize learning in action.[3] Games, simulations, and other digital tools help learners understand phenomena by working with them from the start in complex situations rather than by first mastering isolated facts and skills and later assembling these conceptual building blocks to solve more elaborate, more complete, more realistic, and more sophisticated problems (Gee 2003, 2005; Shaffer 2005a). In such environments, mastery of basic facts and skills are not an effective measure of expertise. Therefore assessments of digital learning need to focus on performance in context rather than on tests of abstracted and isolated skills and knowledge (Delandshere 2002; Hickey et al. 2006; Mislevy 2006; Shute et al., in press).

Fortunately, digital learning environments also provide the potential to assess performance in context, because digital tools make it possible to record rich streams of data about learning in progress. Hardware and software can record *actions* students take in a digital medium as well as *interaction* between learners and between learners and mentors. These data can be recorded dialog and chat, reflective conversations, notebooks, and intermediate work products. Steps have been taken in this direction with simulation-based performance assessments in professional training, certification, and licensure (See, for example, http://www.stisimdrive.com/, http://www.immersion.com/medical/products/endoscopy/bronch/studies.php, http://www.virtualaviation.co.uk/training.html. Also so Bejar and Braun, 1999; Clauser, Margolis, Clyman, and Ross 1997). But these tools are expensive to develop and use, are targeted at very specialized forms of learning, and do not have an established methodology that would make them readily adaptable to the assessment of learning in games and other digital learning environments.

So what assessment methods will use the rich data from digital tools to measure the mastery of complex problem solving that takes place in digital learning?

Here we argue that one way to address this challenge is through an approach to assessment known as *evidence-centered design* (Mislevy 1996, 2006; Mislevy and Steinberg 2003; Mislevy, Steinberg, and Almond, 2001; Shute and Zapata-Rivera, in press). Evidence-centered design is a framework for developing assessments by systematically linking models of understanding, observable actions, and evaluation rubrics to provide evidence of learning. Central to the concept of evidence-centered design—and to our argument here—is the idea of *alignment* between learning theory and assessment method, between evidence and hypothesized mechanisms of thinking and learning in a given arena.

In what follows, we examine how evidence-centered design can address the challenge of assessment in new media learning environments by presenting one specific theory-based approach to digital learning, known as *epistemic games* (Shaffer 2005b, 2007; Shaffer, Squire, Halverson, and Gee 2005a), and describing a method, *epistemic network analysis* (Nash and Shaffer 2008; Nulty and Shaffer 2008), to assess learner performance (and thus also to assess epistemic games[4]) based on this learning theory. The learning theory is supported by empirical studies, and in early trials the assessment method we describe has produced some useful results in analyzing learning outcomes (Beckett and Shaffer 2004, 2005; Hatfield and Shaffer 2005, 2006; Nash and Shaffer 2008; Nulty and Shaffer 2008; Shaffer 1997a, 2002a, 2002b, 2004a, 2004c, 2004d, 2005a, 2005d, 2005e, 2007, in press-a; Shaffer and Squire 2005, 2006; G. N. Svarovsky and Shaffer 2007) However, our goal here is not to argue for the validity of epistemic games or to promote epistemic network analysis. Rather, our hope is to use the theory and its related assessment method to illustrate the concept of a *digital learning system*: a system composed of a theory of learning and its accompanying method of assessment, linked into an evidence-based, digital intervention.

We begin by describing epistemic games,[5] and thus the theory of digital learning at the center of this particular digital learning system. We then explain the mechanics of epistemic network analysis, and provide examples of its use as an assessment. We use the framework of evidence-centered design to show how this approach to assessment can provide a design framework for developing digital learning systems, and conclude with a discussion of the broader implications of a digital learning system based on epistemic network analysis for the development of rigorous assessment of digital learning more broadly.

Learning for a New Century

The problem is by now well known. Technology lets companies send any job overseas that can be done by a skilled worker according to some well-established process (Antráas, Garicano, Rossi-Hansberg, and National Bureau of Economic Research, 2005; Blunden 2004; Burgess and Connell 2006, 2005; Friedman 2005; Hagel and Brown 2005; Hunter 2006; Kanter 2001; Kehal and Singh 2006; Markusen 2005). As a result, young people today need to learn to deal with problems that do not have rote answers. They need to develop judgment and discretion, creative thinking, collaboration, and complex problem solving. But the technology that creates this problem is also part of the solution.

Computer simulations make it possible to create virtual worlds—"environments where people can explore and learn from what they receive back from the computer in return for their exploration" (2002). The power of virtual worlds for learning has been shown in programming environments (diSessa 2000; Harel and Papert 1991; Papert 1980; Resnick 1994), direct manipulation environments (Goldenberg and Cuoco 1998; Serra 1997; Shaffer 1997a, 1997b, 2002b), and game worlds (Adams 1998; Barab et al. 2001; Kafai 1996; Starr 1994). In these worlds, students can learn by tackling problems that are realistic, complex, meaningful, and motivating. And because contemporary computer and video games are profoundly engaging and motivating to young people, they have great potential to teach the kind of thinking that young people need in the digital age (Gee 2003). [AQ3]

Games can help players develop deep understanding of important academic content and valuable forms of thinking from activities that are personally meaningful, experiential, social, and epistemological all at the same time (Shaffer, 2005a, 2007). One way to do this is with epistemic games.[6]

Epistemic Games

In epistemic games, players inhabit a game world in which they are novices training to be professionals of a particular kind: engineers, urban planners, science journalists, and so on. However, simply turning students loose in such a virtual world—no matter how cleverly designed—is a poor instructional strategy because "learners are novices [and] leaving them to float in rich experiences with no support triggers the very real human penchant for finding creative but spurious patterns and generalizations" (2006; 2005b) A key component in turning activity in a virtual world into understanding in the real world is *reflection*: a player's ability to step back from what he or she is doing and talk with peers and mentors about what worked, what didn't work, and why. Studies of expertise, particularly in professional domains, show that the *combination* of action and reflection develops sophisticated ways of thinking (Bruner 1996; Dreyfus and Dreyfus 1986; Kaput 1992; Noss and Hoyles 1996; Schon 1983, 1985, 1987[7]; Shaffer 2004c, 2005c, 2005e, 2007, in press-c) Thus, because direct mentoring by experts is part of any training for expertise, explicit guidance by mentors is part of any epistemic game. [AQ4]

In epistemic games, this guided reflection is based on the kind of mentoring that professionals get in their practicum experiences, rather than the traditional direct instruction of school-based learning or the rote memorization and drill of basic facts and skills that many educational games currently provide. Epistemic games use authentic professional training practices as a model for integrating action and reflection to develop innovative and creative ways of thinking.

For example, in the epistemic game *Urban Science* (http://epistemicgames.org/eg/?cat=14), players become urban planners to redesign their city (Beckett and Shaffer 2005; Shaffer, 2007). They use a geographic information system (GIS—http://en.wikipedia.org/wiki/Geographic_information_system) model to propose land use changes (turning a parking lot into a neighborhood park and playground, or rezoning commercial lots for mixed commercial and residential use) to improve quality of life. In the game they perform the kinds of *actions* that urban planners do in their training: They receive materials that urban planners use, such as a city budget plan and letters from concerned citizens providing information about revenue, pollution, waste, housing and other issues. They conduct a site visit and interview virtual stakeholders. They use a GIS model to create preference surveys and construct proposals for redevelopment. Players also have opportunities for *interaction* with peers and mentors. Real people playing the roles of planning consultants—more generically referred to as *game mentors* in epistemic games—engage players in the kind of reflection that planners use to turn planning activity into understanding of the planning process and the concepts relevant to the planning profession.

Over a decade of research shows that in epistemic games such as *Urban Science*, professional work and reflection on that work with peers and mentors in game form can be a powerful educational tool. Players can learn valuable concepts, principles, practices, and ways of thinking by learning to solve real problems the way professionals do (Beckett and Shaffer 2004, 2005; Hatfield and Shaffer 2006; Shaffer, 1997a, 2002a, 2002b, 2004a, 2004c, 2005a, 2005d, 2005e, 2007, in press-a). In the epistemic game *Digital Zoo* (http://epistemicgames.org/eg/?cat=15), for example, players learn physics and engineering by working as biomechanical engineers to help design characters of the kind seen in computer-generated animation films like *A Bug's Life*[8] (Shaffer 2007; G. N. Svarovsky and Shaffer 2004, 2006a, 2006b, 2006c, 2007). In the epistemic game *science.net* (http://epistemicgames.org/eg/?cat=10), players learn about ecology, genetics, communications technologies, and other current issues by working as science journalists to create an online science newsmagazine (Hatfield and Shaffer 2006; Shaffer 2007; Shaffer and Squire 2006)

Epistemic Frames

Epistemic games are based on a specific theory of learning: the *epistemic frame hypothesis* (Shaffer 2004a, 2005a, in press-a, in press-d). The epistemic frame hypothesis suggests that any community of practice has a culture (Rohde and Shaffer 2004; Shaffer 2004a, in press-a, in press-d) and that culture has a grammar, a structure composed of:

- *Skills:* the things that people within the community do
- *Knowledge:* the understandings that people in the community share
- *Identity:* the way that members of the community see themselves
- *Values:* the beliefs that members of the community hold
- *Epistemology:* the warrants that justify actions or claims as legitimate within the community

This collection of skills, knowledge, identity, values, and epistemology forms the *epistemic frame* of the community. The epistemic frame hypothesis claims that: (a) an epistemic frame binds together the skills, knowledge, values, identity, and epistemology that one takes on as a member of a community of practice; (b) such a frame is internalized through the training and induction processes by which an individual becomes a member of a community; and (c) once internalized, the epistemic frame of a community is used when an individual approaches a situation from the point of view (or in the role) of a member of a community (Shaffer 2004a, 2004b).

Put in more concrete terms, engineers act like engineers, identify themselves as engineers, are interested in engineering, and know about physics, biomechanics, chemistry, and other technical fields. These skills, affiliations, habits, and understandings are made possible by looking at the world in a particular way: by *thinking* like an engineer. The same is true for biologists but for different ways of thinking—and for mathematicians, computer scientists, science journalists, and so on, each with a different epistemic frame.

Epistemic games are thus based on a theory of learning that looks not at isolated skills and knowledge, but at the way skills and knowledge are systematically linked to one another—and to the values, identity, and ways of making decisions and justifying actions of some community of practice.

The key step in developing the epistemic frame of most communities of innovation is some form of *professional practicum* (Schon 1983, 1987). 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.

Epistemic games are simulations of how professional practica develop the epistemic frame of a professional practice: They recreate the *participant structures*[9] (the forms of action and of interaction) of a practicum that provide systematic opportunities for participating in a practice and reflecting about that practice (Shaffer, in press-b; Shaffer, Squire, Halverson, and Gee 2005b). The activity structure of the game can thus be represented in terms of a *frameboard*, a version of the game storyboard that specifies explicitly the actions taken by players and the forms of reflection expected in each participant structure of the game over time. The frameboard also specifies for each participant structure (a) the expected elements of the epistemic frame that will be developed and/or linked at that point in the game and (b) the expected evidence of such development or linkage.[10]

In other words, epistemic games use the epistemic frame hypothesis to translate the activities of a practicum experience for professionals into a digital learning game, and the frameboard links the resulting game activities to the desired outcomes of epistemic frame development. In this sort of learning we do not sacrifice the mastery of facts and skills; rather that mastery is attained in the context of motivating goals and activities—and, thus, learned in a deeper, more meaningful, and more enduring way. In this sense, epistemic games provide a theory-based approach for turning activity in a virtual world into the epistemic frames of professional understanding in the real world.

Assessing Epistemic Frames

Evidence-centered design suggests that any system of assessment begins with a conception of the capabilities one wants to develop in a learner; from there develops the machinery of the assessment system, such as tasks, rubrics, and scoring models (Messick 1994). That is, we start with the things we want students to learn and then identify the actions they might take that will provide evidence of mastery. A different conception of knowledge requires different assessment settings and different actions that will provide evidence of learning—different ways of capturing, expressing, representing, and reporting students' capabilities. And critically in evidence-centered design, the forms of assessment follow from the underlying theory of learning.

To assess epistemic games, then, we begin with the concept of an epistemic frame. The kinds of professional understanding that such games develop is not merely a collection of skills and knowledge—or even of skills, knowledge, identities, values, and epistemologies. The power of an epistemic frame is in the connections among it constituent parts. It is a network of relationships: conceptual, practical, moral, personal, and epistemological.

Power to the Party

There is a body of work on using concept maps (http://en.wikipedia.org/wiki/Concept_map) to assess knowledge and belief structures (Carley and Palmquist 1992; Shute and Zapata-Rivera, in press). Social network analysis (http://en.wikipedia.org/wiki/Social_network)[11] provides a robust set of analytical tools for representing such networks of relationships, including complex and dynamic relationships of the kind that characterize epistemic frames (Brandes and Erlebach 2005; Wasserman and Faust 1994). Of course, social network analysis was developed to provide insight into relationships among and between individuals and groups, rather than relationships within the conceptual, practical, moral, and epistemological world of an individual. Thus, epistemic network analysis adapts the tools of social network analysis to a different—albeit related—domain.

Using social network analysis we might examine the relationships among a group of people meeting for the first time at a cocktail party.[12] To do so, we might take a photograph of the party at appropriate intervals—perhaps every time the music changes, every time someone orders a drink, or at a fixed time interval, depending on the nature of the party and our hypotheses about the people and relationships involved and the social forces at work. If we make the assumption that people who spend more time in the same conversational group develop a closer relationship over time, we can quantify the social network being developed at the party by summing, for each pair of partygoers, the number of times they are recorded in the same conversational group during the party. Once quantified in this way, social network analysis provides a wide range of analytical tools for investigating the properties and processes at work in the social relationships of the party.

If we think of the "party" not as a collection of individuals in a room, but rather as a collection of elements from the epistemic frame of one player in an epistemic game, then we can use the same analytical tools to conduct epistemic network analysis.

Epistemic Network Analysis by Analogy

Epistemic network analysis is thus a form of network analysis for assessing epistemic frames. By analogy to the social network analysis of a cocktail party, epistemic network analysis looks at an "epistemic party" in which elements of the epistemic frame are in use by a particular player (and therefore "in conversation" with one another) over time. Although the frame elements do not have the same kind of agency as partygoers, the state of their relationships can be usefully modeled in a similar fashion.

In formal terms, the epistemic frame of a given profession, P, has elements $f_{1 \dots n}$, where each f_1 is some skill, knowledge, value, or form of identity or epistemology that is part of the profession P. In this sense each f_1 is a partygoer in this epistemic party. The epistemic game based on P can be described as a series of activities about which we collect such data, D_t, which represents information about the players during the activities at time t. In this sense, each D_t is the picture of the epistemic party taken at time t.[13]

For any participant p, we can look at $D^p{}_t$, a subset of D_t containing the evidence that at time t that player p is using one or more of the elements of the epistemic frame of the profession P. A player with a more robustly developed epistemic frame will tend to say and do things in the game situations that reflect this richer structure—which in turn reflects developing expertise. In this sense, $D^p{}_t$ is one of the "conversational groups" of the epistemic party, showing how player p is using elements of the epistemic frame at time t. And for a given player, p, the sequence of data $D^p{}_{1 \dots e}$ spanning from the beginning ($t = 1$) to the end ($t = e$) of the game represents the *play history* of player p's participation in the game.

Data

By way of example, here and throughout this section, we draw on data from two epistemic games, *Digital Zoo* and *Urban Science*, both of which are described briefly in the previous section of this article. In each case, we look at data from approximately 80 hours of game play conducted over four weeks (approximately four hours per day) during a summer enrichment program. In each case, data was collected in two forms. First, we collected work done by each player: a set of notebooks, design documents, reports, and other work products produced during game play. Second, we recorded mentor interactions with each player: conversations in individual and group meetings with mentors during the game. These records of work and interaction were then assembled into a play history for each player, and coded using the frame elements identified in the frameboard for the game (Beckett and Shaffer 2005; Nash and Shaffer 2008; Nulty and Shaffer 2008; G. N. Svarovsky and Shaffer 2007).

Network Graphs

Let us consider, for example, an excerpt from a player's design notebook in *Digital Zoo*. In this excerpt, we see the player use several elements of the epistemic frame of engineering. The excerpt could be coded for the skill of comparing design alternatives (S\CA), for the values of designing to meet a client's needs (V\CN) and producing reliable designs (V\RD), and for the epistemology of making judgments based on quantifiable tests of performance (E\QT). The episode would almost certainly not be coded for other elements of the epistemic frame of engineering used in the game, such as knowledge of the center of mass (K\CM), or the identity of seeing oneself as an engineer (I/SE), for which there is no evidence in this data excerpt:

In other words we have data to suggest that at this point in the game this player was linking a particular engineering skill, two different professional values, and a way of justifying action into a unified performance of the practice of engineering—or more precisely, a legitimate but perhaps not-yet-expert performance of the practice.[14] In terms of the analogy to a cocktail party, at the epistemic party for this player at this point in time partygoers S\CA, V\CN, V\RD, and E\QT are talking to one another. K\CM and I\SE have either not yet arrived or are checking their voicemail in another room, outside of this particular picture frame.

This allows us to construct a network graph $G^p{}_t$ showing the relationships among the frame elements in use in $D^p{}_t$—that is, in the data we have about player p at time t:

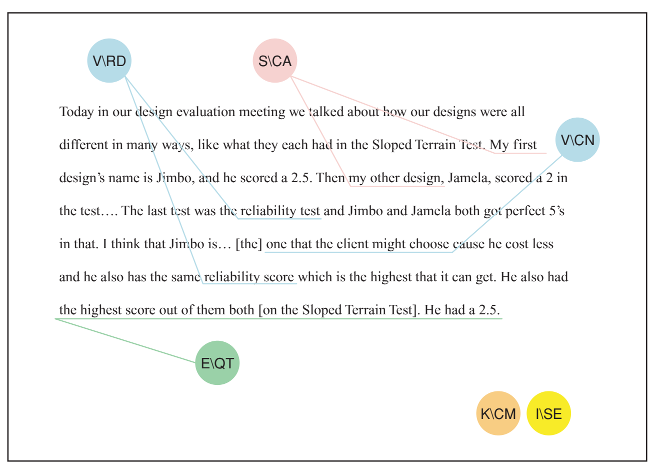

Figure 1 Coded data excerpt from a play history. [AQ5]

In this Kamada-Kawai spring-mass model (https://nwb.slis.indiana.edu/community/?n=VisualizeData.Kamada-Kawaii) of the epistemic network, elements of the player's epistemic frame are shown as *nodes* (circles) in the graph (Kamada and Kawai 1989). Frame elements that are linked in the data $D^p{}_t$ are shown as nodes connected by *arcs* (the solid lines) in the diagrams.[15] Nodes that are not linked in $D^p{}_t$ are shown without connecting arcs.

We note that this is only a partial representation of this player's network—or, more precisely, only a part of the network graph that represents data we have about this player's epistemic frame. There are other elements of the epistemic frame of engineering at play in *Digital Zoo*, and we present this partial view for illustrative purposes here.

Summing Strips over Time

In technical terms, a network graph such as this represents the epistemic frame in use by player *p* in the strip of time *t* based on the evidence in our data set *D*. The term *strip of time* comes from Goffman's (1974) work on *frame analysis* (http://www.ccsr.ac.uk/methods/publications/frameanalysis/), and with it the idea of a *frame*: a set of "principles of organization which govern events" (p. 10). In frame analysis, Goffman describes action in terms of strips, where a strip is a "slice or cut from the stream of ongoing activity… as seen from the perspective of those subjectively involved in sustaining an interest in them" (p. 10). Strips of activity are segments or units into which ongoing activities are divided for the purposes of analysis.[16] The photograph of the cocktail party at time *t* is thus a picture of the party during some strip of time, just as this network graph is a representation of the epistemic party at time *t*.

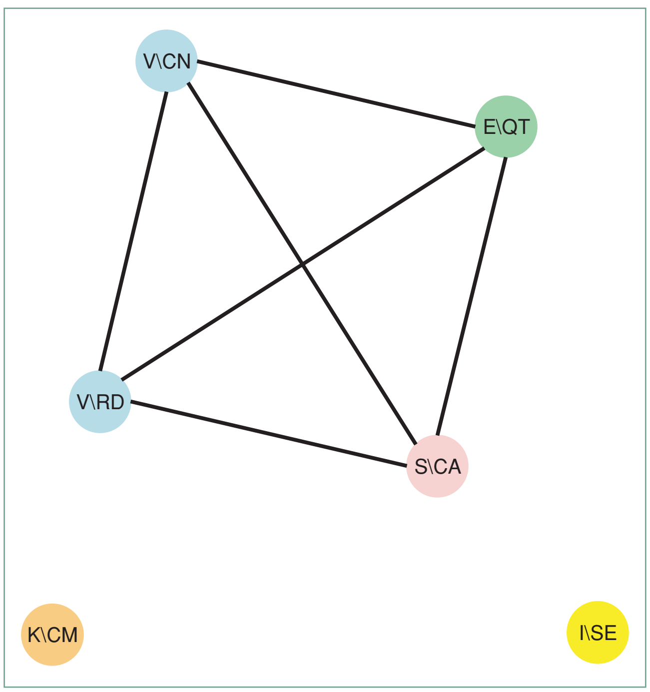

Figure 2 Network graph of frame elements at one point in a play history.

In the cocktail party example we could quantify the social network of the partygoers by summing, for each pair of partygoers, the number of times they are recorded in the same conversational group during the party. Similarly, in this epistemic party we can quantify the epistemic network for player *p* by summing, for each pair of frame elements, the number of times they are recorded in the same strip of activity during an epistemic game.

We can then create a cumulative network graph, where frame elements (nodes) that are linked more often in the data are closer to each other than those that are linked less often in the data. That is, the length of the arc between two frame elements in the network graph is inversely proportional to the number of strips of activity in which they co-occur, with the distance computed, as before, using a Kawada-Kamai spring-mass model. Here, for example, is the final epistemic network graph showing one player's epistemic frame for engineering after 80 hours of playing *Digital Zoo*:

Notice a dense core of skills, knowledge, and values (red, orange, and blue nodes) at the center of the network graph (which is toward the right of the

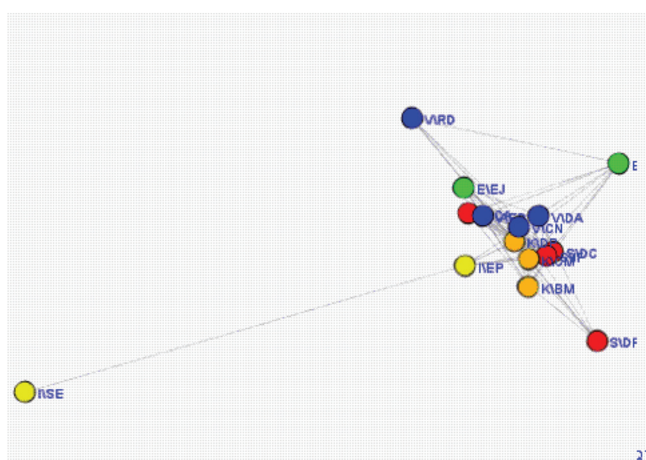

Figure 3 Final network graph[17] for one player in *Digital Zoo*.

image above). At a greater distance from this core of skills, knowledge, and values are elements of identity (yellow nodes) and epistemology (green nodes)—as well as additional skills and values—that are progressively less central in the epistemic frame of this particular player. Furthest away is a relatively isolated frame element: the identification of self as an engineer. Here is a player who clearly mobilizes many of the attributes of an engineer in solving complex problems, but rarely makes explicit reference to herself as an engineer in that process.

From Strip to Slice, and Everything Nice

We can use the same technique to produce a graph for a player not only at the end of the game, but at any point of time in the game. Instead of summing all of the strips of activity in the game, we sum the strips up to some particular time.[18] In this sense, the trajectory of development of an epistemic frame can be mapped as a dynamic network graph, or series of *slices* over time, where each slice shows the state of the players' epistemic frame at a different point in time.

For example, here are three slices of frame development for this same player of *Digital Zoo*:

The slices these graphs represent come from the early, middle, and late parts of the game (moving from left to right in the figure above). Notice that early in the game, the frame is relatively loose, contains relatively few elements of the epistemic frame of the profession, and is relatively even in the distribution of elements.[19] In the middle of the game, the frame contains more elements, and begins to develop a central core that now includes knowledge and some values, as well as skills. Later in the game, values and epistemology become more central in the player's frame. All of the elements are incorporated, and the core includes still more elements. The network as a whole becomes more dense over time. There are more elements included, and they become more closely connected; that is, there have been more and more slices in which we have observed an interplay among these frame elements.[20]

Weighted Density

We can begin to quantify this change in the linkage of an epistemic network by computing its *weighted density* at each point in time. The weighted density is calculated as the square root of the sum of the squares of the associations between individual elements in the frame. (For details on computing this and other elements of epistemic network analysis, please see the appendix.) As such, it provides a measure of the overall strength of association of the network, emphasizing the dense core of the graph as being central to the strength of the epistemic frame. The figure below shows that, as the selected images suggest, this player's network graph became more tightly linked through game play as measured by the weighted density of her epistemic frame.

We can use the weighted density of an epistemic frame to measure changes in an epistemic frame over time, and to associate those changes with specific elements of game play. For example, in one study of *Digital Zoo* we computed the weighted density of players' epistemic frames as reflected in interview questions about engineering before and after game play (Nulty and Shaffer 2008). We then compared players who reported getting help from mentors in the game with players who did not. Players who reported getting help from mentors showed a significantly greater change in the weighted density of their epistemic frames (see figure below).

Trajectories of Frame Development

Of course, we are interested not only in the overall density of a player's epistemic frame, but also the

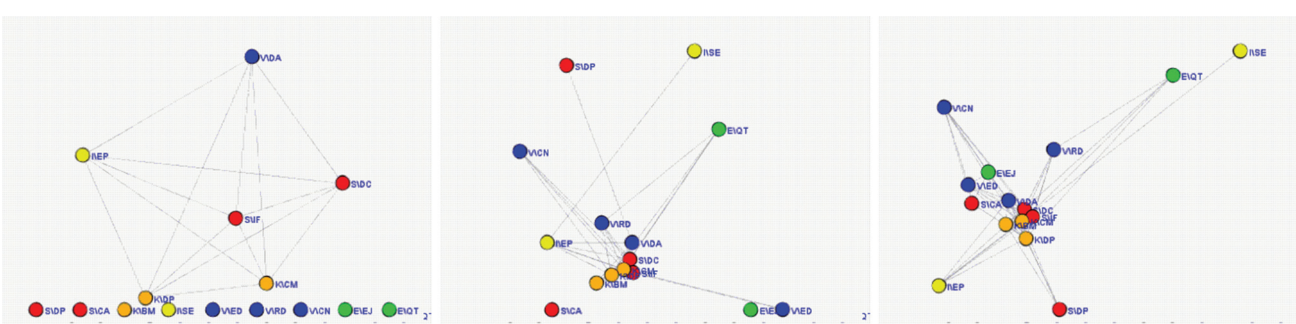

Figure 4 Network graphs from different slices of the play history for one player in *Digital Zoo*.

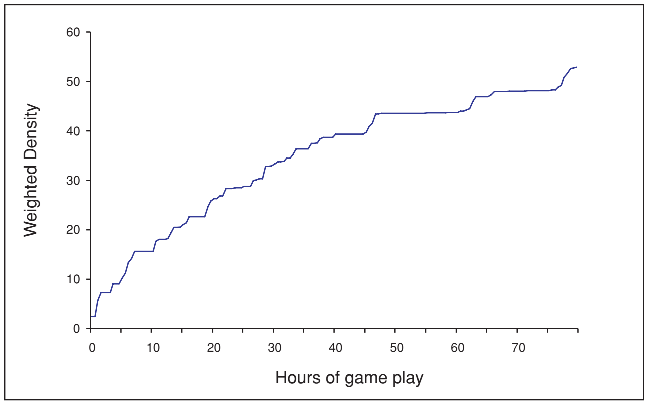

Figure 5 Change in weighted density of one player's epistemic frame over time.

developing shape of his or her frame: which elements of the frame are linked most closely, and therefore are most central in the overall frame. In particular, we want to be able to assess the extent to which a player not only uses elements of the epistemic frame of a practice, but the extent to which he or she uses elements of the frame the way a more experienced practitioner does. That is, we want to assess the extent to which elements of the frame become linked in accord with the ways of doing, knowing, valuing, and thinking of a valued social practice.

Dynamic network graphs of the slices of a developing epistemic frame provide a powerful way to address such questions. For example, in our analysis of the epistemic game *Urban Science*, we compared the cumulative epistemic frame of the game mentors with the cumulative epistemic frame of the players (Nash and Shaffer 2008). That is, we constructed epistemic network graphs representing the collective frame of all of the mentors and all of the players respectively. The figure below shows four slices from each frame, taken at the same point in time during game play.[21] The slices from the mentors' frame are on the left in each pair of images. The slices from the players' frame are on the right in each pair. Moving down the figure, each pair represents the mentors' (left) and players' (right) frame at a progressively later point in time in the game. The topmost image is from the first day of game play. The final image is after three weeks of play. In this analysis we collapsed individual frame elements into larger categories of skill, knowledge, identity, values, and epistemology, which are labeled on the individual network graphs and represented as blue, yellow, green, red, and orange nodes respectively.

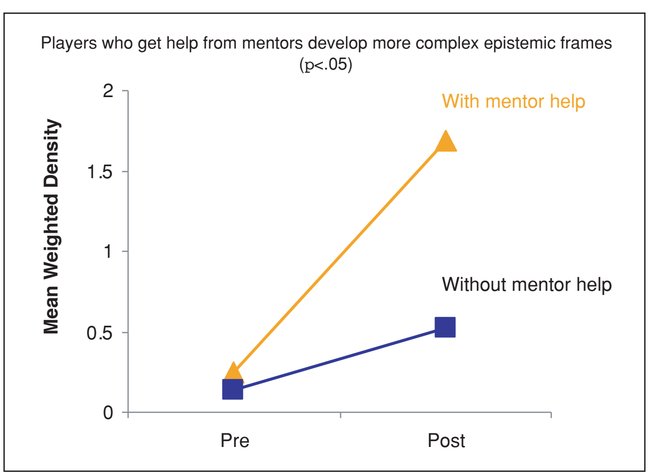

Figure 6 Quantifying the relationship between mentor help and epistemic frame development.

Notice in this set of slices that both the mentors' and the players' frames start out with relatively loose connections, in a similar configuration. At each point in time, the general shape of the players' frame follows the mentors' frame although the elements are not as closely connected. There is, for example, a lag in players' frame evident in the second pair of slices: the node for epistemology is significantly further from the core for the players, and only later moves in to a position similar to the same node in the mentors' frame. Overall, however, these slices suggest that the players are developing a frame that reflects the professional thinking modeled by the mentors in the game.

Relative Centrality

We can quantify the "shape" of an epistemic frame in this sense by computing the *relative weight*, or relative centrality, of each node and comparing them. By extension from the weighted density of the network as a whole, we compute the *weight of a node* from the square root of the sum of squares of its associations with its neighbors. Again, this measures the strength of association for a given node, emphasizing nodes with tighter linkages to individual neighbors. (And, again, for details on computing this and other elements of epistemic network analysis, please see the appendix.)

The weight of a given node will tend to rise as the overall density of the network rises. But we can quantify some of the *shape* of the graph by computing the relative weight of a node—that is, by expressing the weight of a given node as a percentage of the weight of the heaviest node in the network. This provides a measure of the centrality of the node within the network, that is, the extent to which it is (or is not) part of the dense central core of the network.

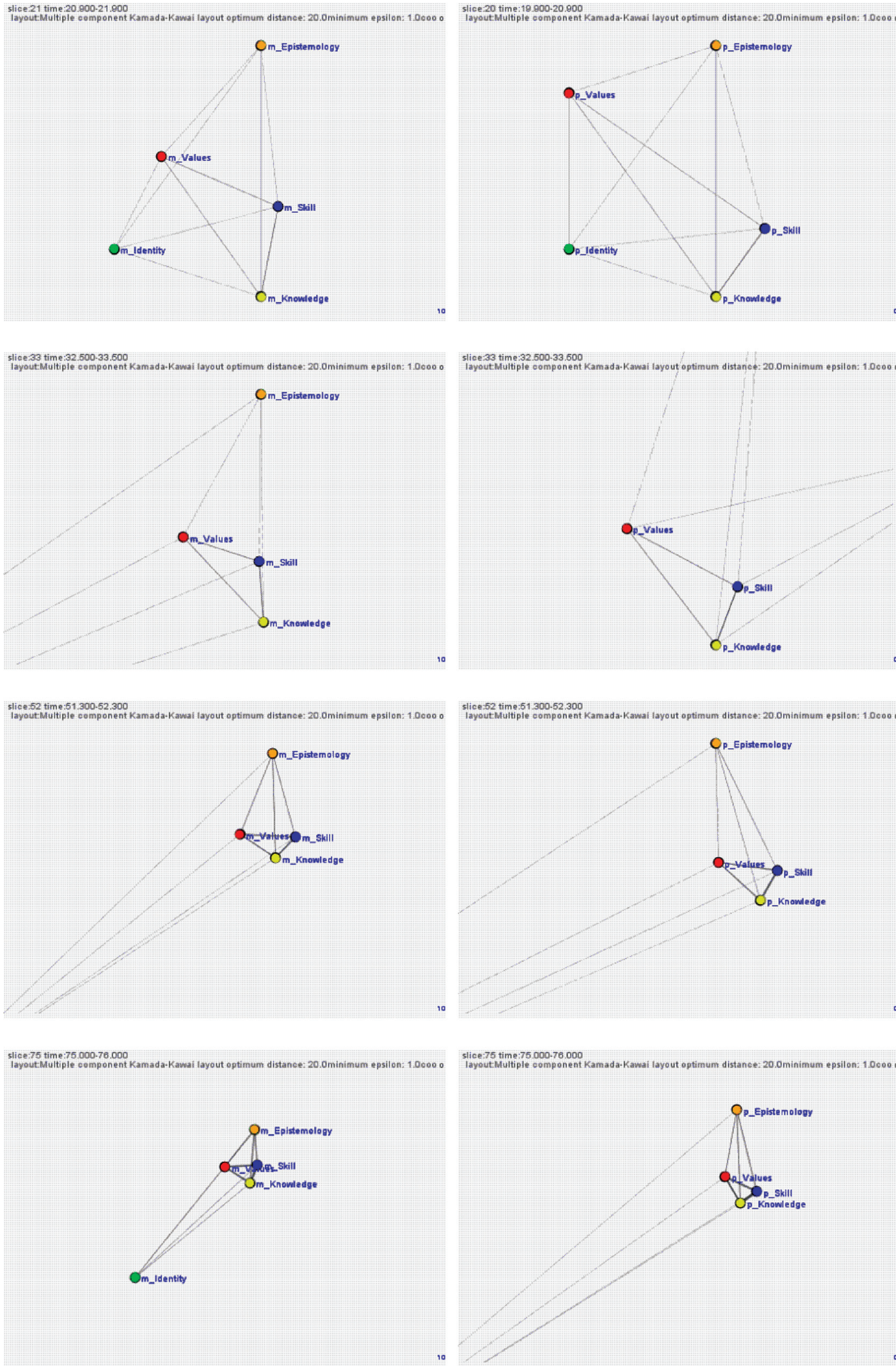

Figure 7 Change in Mentors' (left column) and Players' (right column) frames over time.

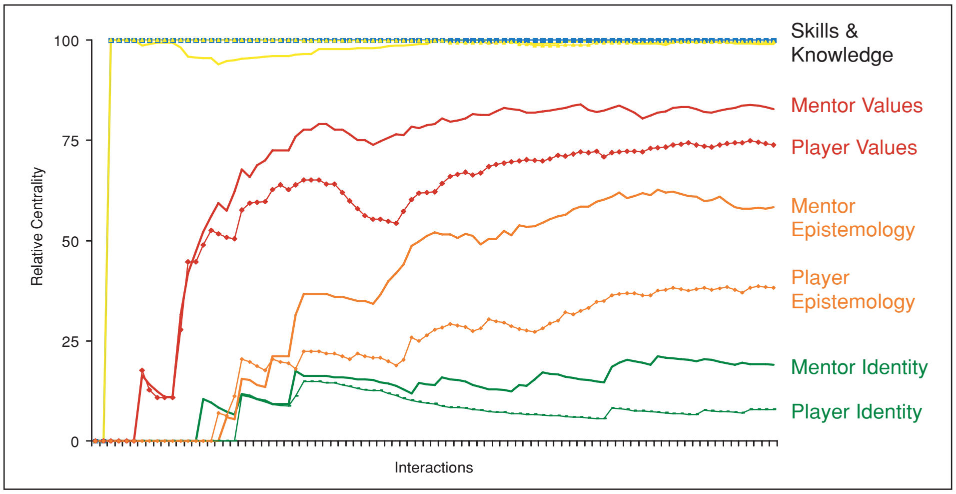

Figure 8 Change in relative centrality of frame elements over time.

The following figure shows a graph of the relative centrality of nodes for the epistemic networks of mentors and players from which the slices in the previous figure were taken (Nash and Shaffer 2008):

Notice that, as the network graphs suggested, the trajectory of frame development for players follows the model provided by the mentors in their interactions: The lines for each frame element showing the relative centrality over time follow a similar shape. The relative centrality of elements in the players' frames are consistently lower, suggesting their frames are not as tightly linked, but the order of centrality of frame elements is the same for both groups.

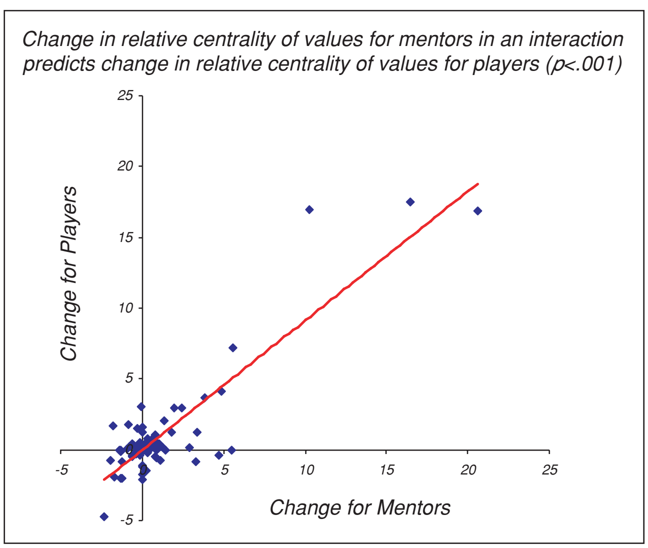

Figure 9 Relationship between change in mentor values and player values over time.

We can thus test the hypothesis that changes in the players' frame is associated with changes in the model frame the mentors are providing through interactions in the game by comparing the changes in the relative centrality of nodes in the two networks over time. And, in fact, in this case, the changes in the centrality of values, epistemology, and identity in the player's frame can be predicted from changes in the mentors' frame ($p < .05$ for each result using a linear regression model). [AQ6]

The scatterplot below shows the regression for change in values:

Notice that (not surprisingly) the data contains several large changes in the relative centrality of values. These correspond, of course, to places in the preceding graph where there are large jumps in the relative centrality of values in the epistemic frame. This suggests that not only are the changes correlated overall, but that key interactions in the game make significant contributions to the development of values in the game.

The power of epistemic network analysis is, of course, precisely that we can test such hypotheses with the data. Each slice of the epistemic network over time is associated with specific events during game play. We can thus explore whether particular kinds of events in the game, perhaps in combination with specific features of a player's epistemic network at that time, lead to significant changes in a player's developing epistemic frame.

That is, we can use epistemic network analysis to assess the development of epistemic frames, and thus development of complex thinking skills through epistemic game play.

Epistemic-Centered Design (ECD)

As we said at the beginning of this paper, our purpose in describing epistemic games and epistemic network analysis was to use them as a worked example of one approach to assessment in the digital age. That is, to show how the relationship between a particular theory of digital learning and its associated assessment method shed light on assessment design in a changing landscape of learning. In what follows, we use the lens of evidence-centered design to examine the connections between epistemic games and epistemic network analysis—and the implications that has for assessment design more broadly.

The Models of Evidence-Centered Design

As Messick (1994) suggests, the core of evidence-centered assessment is the connection between learning, behavior, and setting:

> Begin by asking what complex of knowledge, skills, or other attributes should be assessed, presumably because they are tied to explicit or implicit objectives of instruction or are otherwise valued by society. Next, what behaviors or performances should reveal those constructs, and what tasks or situations should elicit those behaviors? (p. 16).

This quotation certainly fits traditional assessment design, where the view of knowledge is facts and skills, and simple tasks suffice to provide evidence. Evidence-centered assessment design (Mislevy and Riconscente 2006; Mislevy and Steinberg 2003) was developed specifically to design assessments for challenges that lie beyond familiar testing practices—for opportunities provided by advances in technology, such as simulation-based assessment, demands for assessing higher-level capabilities such as model-based reasoning in scientific investigations, and recognition of the situated nature of proficiencies in domains such as classroom teaching.

In its full form, ECD views assessment design in terms of layers. These layers include (a) analysis of the domain on which to ground the assessment; (b) specification of a conceptual assessment framework, or formal model that embodies the assessment argument; and (c) data structures and processes for implementing and delivering assessment. By providing an epistemic frame for assessment that encompasses new as well as familiar forms, evidence-centered design helps designers develop assessments more efficiently and with more explicit arguments to support the validity of the results.

In this section of the article we examine the extent to which epistemic network analysis functions as a conceptual assessment framework for epistemic games.

In evidence-centered design, a conceptual assessment framework focuses on assessment in a way that is strongly connected with a view of learning. It is composed of linked models: a student model, a pair of evidence models, and a set of task models (Mislevy 2006; Mislevy, Steinberg, and Almond 1999).

The *student model* is motivated by the learning theory that underlies the assessment system. It specifies the relevant variables or aspects of learning that we want to assess, at a grain size that suits the purpose of the assessment. As many of the characteristics of learning that we want to assess are not directly observable (a problem that has plagued psychology in general and cognitive science in particular for decades), the student model provides a probabilistic or proxy model for making claims about the state, structure, and development of a more complex underlying system.

To make claims about learning as reflected in the student model, we thus have to develop a pair

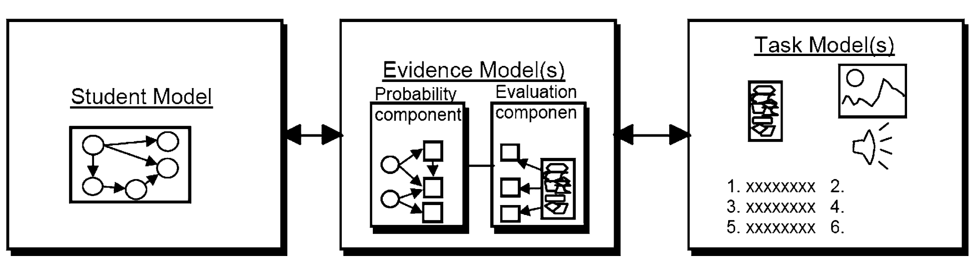

Figure 10 The framework of evidence-centered design (Mislevy 2006).

of *evidence models*. The *evaluation* component of the evidence model specifies "the salient features of whatever the student says, does, or creates in the task situation," as well as the rules for scoring, rating, or otherwise categorizing the salient features of the assessment (Mislevy, Steinberg, and Almond 1999). That is, the evaluation component specifies the things being used for the assessment and the criteria on which they are evaluated. The *probability* component of the evidence model specifies the rules by which the evidence collected in the evaluation is used to make assertions about the student model. Once information is collected and analyzed in the evaluation component, the probability component provides "the machinery for updating beliefs about student model variables in light of this information" (Mislevy, Steinberg, and Almond 1999). Taken together, these evidence models provide a chain of inferential reasoning from observable performance to changes that we believe are significant in an individual's cognitive, social, emotional, moral, or other forms of development.

Finally, the *task model* provides a set of "specifications for the environment in which the student will say, do, or produce something" (Mislevy, Steinberg, and Almond 1999). It is a description of the assessment environment, including the resources available and specifications for the work product that represents what the learner does, makes, or says as input to the evidence model. That is, the task model specifies the conditions and forms under which data are collected.

In describing these models, Mislevy points out that the "data" being collected are not restricted to traditional formal, structured, pencil and paper assessments. Data that flows from a task model can include information about the context, the learner's actions, and the learner's past history or particular relation to the setting (Mislevy 2006).

The goals of evidence-centered design are not to prescribe (or proscribe) particular forms of evidence or theories of learning. Rather, the aim is to provide a theoretical mechanism for making the process and artifacts of assessment formal and explicit—to describe a method for structuring the collection, management, and interpretation of information, making it possible to use the same argument structure across students to provide evidence for claims about learning. Evidence-centered design argues for the importance of linking learning theory and assessment—and provides the conceptual machinery for accomplishing that vital task.

Epistemic Network Analysis as ECD

Evidence-centered design thus provides both a theory and a method for linking epistemic games and epistemic network analysis. And contrariwise, epistemic games and epistemic network analysis provide an example of how and why evidence-centered design is a viable framework for thinking about the assessment in the digital age.

From the point of view of evidence-centered design, the epistemic frame hypothesis guides the development of a student model based on the elements of the epistemic frame of a profession or other socially valued practice as developed in some practicum experience. The activities of the practicum—and the hypothesized relationship of those activities to the development of an epistemic frame—are translated into a frameboard that serves as a task model to guide the development of specific game activities and data collection apparatus.[22] The nature and connectivity of a student's epistemic network are the target of assessment. The student model itself consists of variables that represent key features of an epistemic network, such as network density and relatively centrality of nodes.

The evidence model is composed of two parts: a set of play histories that are coded based on the elements of the professional frame, and a set of analytic tools for translating the frame elements as they occur in the play history into an epistemic network representation. In this way, epistemic network analysis lets us re-represent the strips of activity in the

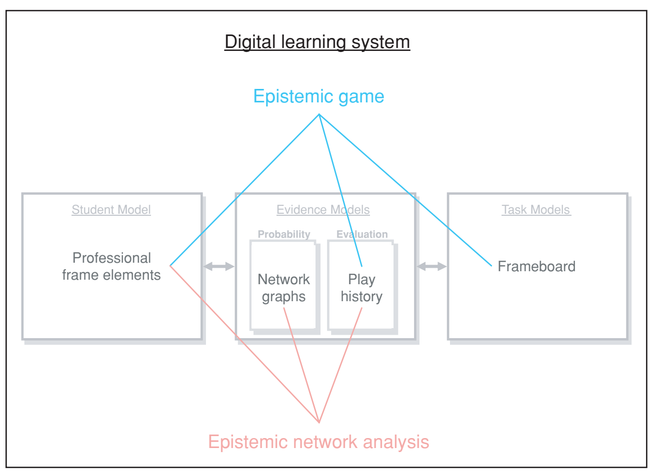

Figure 11 Epistemic games and epistemic network analysis as a digital learning system.

game as slices in the trajectory of development of players' epistemic frames. These slices can then be quantified and analyzed so as to assess individual frame development, and also associate particular game activities with that development.

As the figure below suggests, evidence-centered design shows how epistemic games and epistemic network analysis align and integrate learning activities and assessment strategies:

That is, epistemic games and epistemic network analysis are an example of a digital learning system that exemplifies key principles of evidence-centered design as a framework for assessing complex problem solving.

Discussion: A Digital Learning System

What, then, do epistemic games and epistemic network analysis—as a retrospective example of evidence-centered design—tell us about digital learning systems, and thus about assessment in the digital age?

Alignments

Central to the digital learning system of epistemic games and epistemic network analysis is a series of linkages or *alignments* that are key features of learning in the digital age.

Activity and assessment. In domains that require complex problem solving rather than mastery of isolated facts and skills, assessment needs to take place *in situ*. That is, we need to assess understanding in the complex conditions of practice in which learning takes place. The mechanisms of epistemic network analysis make this possible in epistemic games by providing a task model that combines both assessment and learning. The materials of game play become the objects upon which assessment is based. Learning and assessment are aligned at the level of play itself.

Formative and summative assessments. As a result of this linkage between learning and assessment, epistemic games and epistemic network analysis provide a second type of important alignment: a common platform and language for assessing both the learning environment and the individual development within that environment. Because assessment using epistemic network analysis is based on—and tied to—activities in the game, data about the impact of specific activities arises directly from an analysis of learner performance during play. This means that the system as a whole can be adapted to individual learners—and feedback from assessment is both summative (it describes progress along a desired continuum) and formative (it provides information to guide subsequent pedagogical choices).

Learning outcomes and real-world outcomes. Perhaps most important, though, epistemic games and epistemic network analysis align learning outcomes inside the game with outcomes that are valuable in the world outside of game play. The target frames of assessment are based on ways of thinking exhibited in professions and other socially valued practices. We can validate the target frames of epistemic network analysis in game play through transfer studies or other longitudinal data collection—that is, we can see whether frame development in the game is associated with better performance on some test, or in long-term career choices or other forms of success in the world. But we can also validate epistemic network assessments by showing that performance of game players follows the same pattern and trajectory of work as real world practitioners.

Measurement Questions

Epistemic network analysis shows promise as a basis for the assessment of epistemic games and perhaps digital learning environments more broadly. The significant difference between the graphs of players with and without mentors (figure 6), for example, and the level differences between players' and mentors' centrality values (figure 8) suggest that indices from epistemic network analysis can be useful for group comparisons and experimental studies of interventions. But whether, and how, epistemic network analysis can be further extended to the assessment of epistemic networks of individual students will require further examination of several key measurement issues:

Sorting out learning and accumulating evidence. The evolution of the epistemic network graph over the course of a game is affected by several factors. One is that the graph itself is not equivalent to a player's epistemic network, but a reflection of actions that are made through the player's epistemic network at each given point in time. The evolution of the graph depends partly on how long the game has been played so far, the nature of the game conditions the player experiences (some situations may be more likely to evoke statements of values, for example, or identities), and the changing nature of the player's actual epistemic network as it develops through these experiences. Further research can sort out some of these factors, for example through comparisons of beginners' and mentors' epistemic network graphs in a given game. Epistemic network graphs for both evolve, but as figure 8 shows, the mentors' graphs show more rapid development and higher ultimate levels of density for values and identity, presumably because their actual epistemic networks were more richly interconnected from the start. Using mentors' epistemic network graphs as baselines may enable us to better discern effects of learning from effects of accumulating data.

Generalizability. How tightly is what we learn from an epistemic network developed in a given game bound to the specifics of that game and that domain? We need to determine the extent to which things we learn about a player's epistemic frame in a game like *Digital Zoo* are relevant to his or her values, identities, and epistemological proclivities beyond that context. And we need to establish that epistemic network analysis can provide inferences that are in some way commensurable across different contexts.

Focused task design. As we noted in the discussion of figure 9, some time slices seem to be particularly good at evoking evidence about certain aspects of an epistemic frame. What are the features of these situations? Identifying them will provide valuable information both for tuning games for instruction and for developing focused episodes in a game for assessment. As we look ahead to measurement models, we may be able to take advantage of developments in psychometrics for designing tasks and modeling performances that exploit theory about how people acquire and use knowledge in context.

Statistical modeling. Assessing individuals usually requires some indication of the accuracy of inferences. In traditional psychometrics, classical test theory provides reliability indices and models such as item response theory provide measurement errors for any reported estimates of student model variables. We need to develop empirical methods, such as Tukey's jackknife (http://en.wikipedia.org/wiki/Bootstrapping_(statistics); Mosteller and Tukey 1977) to compare the variation of estimates computed with different samples of time slices. [AQ7]

New Models

Our point, of course, is not that epistemic games and epistemic network analysis are the only or even the best solution to the challenges of education in the digital age. For example, Bogost, extending Papert's work, suggests that we conceptualize digital learning in terms of *procedural literacy* (Bogost 2005; Papert 1980). Similarly, Csíkszentmihályi's work on flow provides a powerful framework for thinking about the impacts of activity in digital environments (Csikszentmihalyi, 1996). Such formulations may be compatible with the notion of epistemic frames—not least because work on epistemic games is similarly based on some of Papert's ideas—but our intent here is not to compare and contrast different theories of digital learning. Similarly, the concept of evidence-centered design may have applications to good game design more generally, but such an exploration is beyond the scope of the discussion.

Our goal, rather, has been to offer one particular learning theory and approach to assessment as a worked example of one way to use evidence-centered design in digital learning environments—and thus to exemplify some of the key issues that any digital learning system will face.

The digital learning system we describe here shows a practical, theory-based approach to assessment that is fundamentally about *connections between* skills, knowledge, identity, values, and ways of thinking. It integrates new forms of learning with new modes of assessment into a coherent system. Epistemic games and epistemic network analysis provide a method for developing competence in the complex thinking of real world professionals, and for assessing that development in context and in use. They thus provide a mechanism to support sociocultural learning, and a means to provide evidence for the kinds of situated understanding that results.

It is plausible that the concepts and methods from the epistemic frame hypothesis and epistemic network analysis could be applied to other learning environments that similarly emphasize the development of understanding through situated action, and that take as a premise the idea that facts and skills are only meaningful when linked in a larger network of practice. But whatever tools of learning and assessment other environments use, the analysis here suggests that they will need to be integrated into a coherent system that aligns learning, assessment, and real-world outcomes—and that evidence-centered design is an effective framework for thinking about the alignment of learning and assessment in the digital age.

Appendix: The Mathematics of Epistemic Network Analysis

In this section we provide an overview of the mathematics behind epistemic network analysis.

Epistemic Adjacency Matrices

We begin by recalling that the epistemic frame of a given profession, P, has elements $f_{1 \ldots n}$, where each f_1 is some element of the epistemic frame of P. Further, the epistemic game based on P can be described as a series of activities about which we collect data, D.

For any participant p, we can look at $D^p{}_t$, containing the evidence that at time t player p is using one or more of the elements of the epistemic frame of the profession P. To construct an epistemic network from data such as this, we create an adjacency matrix, $A^{p,t}$, for player p at time t, recording the links between elements of the frame for which there is evidence in $D^p{}_t$:

$$(1)\ A^{p,t}{}_{i,j} = 1 \text{ if } f_i \text{ and } f_j \text{ are both in } D^p{}_t$$

We could, of course, use $A^{p,t}{}_{i,j} = L(f_i, f_j)$ for any other function L that measured not just the co-presence of f_i and f_j in the information about player p at time t but the *strength of association* between these elements. This would be the equivalent of assessing the strength of association between two partygoers, rather than only their co-presence in a conversation in a given photograph of the party. For clarity, we describe the simpler case of association here, but the mathematics (and the theory) remain the same regardless of the linking function L used to construct the adjacency matrix.

By representing the epistemic frame in use during a strip of activity as an adjacency matrix, we can use the tools of network analysis to examine the cumulative impact of strips of activity on a developing epistemic frame. We construct the cumulative adjacency matrix for player p, F^p, by summing the adjacency matrices $A^{p,t}$ from time $t = 0$ to the end of the game at time $t = e$ as:

$$(2)\ F^p = \sum_{n=0}^{e} A^{p,n}$$

We can then represent this cumulative adjacency matrix as a network graph, where frame elements (nodes) that are linked more often in the data are closer to each other than those that are linked less often in the data. That is, the length of the arc between f_i and f_j in the network graph is inversely proportional to the value of $F^p{}_{i,j}$, the association between f_i and f_j in the cumulative adjacency matrix, computed using a Kawada-Kamai spring mass model (Kamada and Kawai 1989).

We use the same equation to compute $F^{p,t}$ the epistemic frame for player p at any point of time t in the game. That is, we sum the adjacency matrices for all of the strips up to time t by:

$$(2a)\ F^{p,t} = \sum_{n=0}^{t} A^{p,n}$$

If $A^{p,t}$ represents the epistemic frame in use for player p during strip at time t, then $F^{p,t}$ represents the cumulative epistemic frame for player p at time slice t. We thus represent the development of player p's epistemic frame through a series of cumulative adjacency matrices $F^{p,0} \ldots F^{p,e}$ where e represents the end of the game.

We compute $W(M)$, the weighted density of an epistemic network M (either for a strip of time or a slice of frame development—that is, either for $A^{p,t}$ or $F^{p,t}$) by computing:

$$(3)\ W(M) = \sqrt{\sum_{i,j} \left(\frac{(M_{i,j})^2}{2} \right)}$$

We choose this measure rather than the more traditional network density function because network density measures only the number of links in the network as a percentage of the total links (Brandes and Erlebach 2005; Wasserman and Faust 1994). Such a measure therefore does not take into account the weight of the associations in the network—which is, of course, a more useful measure of the strength of an epistemic frame. Computing weighted density from the squares of the associations weights the core of the network, representing networks with a small number of strong associations as stronger than networks with a large number of weak associations. We divide by 2 because the adjacency matrix is symmetric, thus the formula counts each linkage twice; and we compute the root of the sum of squares to use links as the units of weighted density.

Similarly, we compute the sum-of-squares-centrality $C(f_i)$ or weight of an individual node f_i in matrix M as:

$$(4)\ C(f_i) = \sqrt{\sum_j (M_{i,j})^2}$$

This measure of centrality corresponds well with the position in Kamada-Kawai network graphs—more so than more traditional measures of node centrality such as eigenvalue centrality, betweenness centrality, or k-core centrality—because mature epistemic network graphs (graphs from late in a game or dataset) tend to be dense (most nodes are linked to one another) and have a single core (they do not have distinct subgroups; Koschützki et al. 2005). As a result, there is a strong correlation between the weight of a node as measured by the strength of association to its neighbors and its position in the network. We use root sum of squares both to emphasize strong associations and to keep node weight and weighted density of the network compatible measures.

Finally, we compute the relative centrality $R(f_i)$ of an individual node f_i in matrix M as:

$$(5)\quad R(f_i) = \frac{C(f_i)}{C_{\max}(M)}$$

where $C_{\max}(M)$ is maximum node weight of any node in M.

Acknowledgement

The research reported in this article was supported in part by a grant from the John D. and Catherine T. Macarthur Foundation, a National Science Foundation Faculty Early Career Development Award (REC-0347000), a Spencer Foundation/National Academy of Education Postdoctoral Fellowship, a grant from the Wisconsin Alumni Research Foundation, the Academic Advanced Distributed Learning CoLaboratory, and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin—Madison. Any opinions, findings, or conclusions expressed in this paper are those of the author(s) and do not necessarily reflect the views of the funding agencies or cooperating institutions.

Notes

1. More on evidence-centered design can be found at http://www.education.umd.edu/EDMS/mislevy/papers/.
2. For more on this topic see Shaffer and Gee (2007).
3. For more on this topic, see Gee (2003). A more general list of classic resources on experiential learning can be found at http://wilderdom.com/experiential/ExperientialReferences.html.
4. For more on the relationship of assessment of individuals and assessment of games, see Shaffer (in press-e).
5. A more comprehensive overview of epistemic games can be found in Shaffer (2007).
6. As with any medium, computer games can also have negative effects—although we note that no studies to date have shown this for epistemic games. Our goal in this paper is not to take a particular position on the relative value of games in general or of epistemic games in particular; rather, we examine the challenges in assessing the kind of positive outcomes that can accrue in such settings.
7. Donald Schon's work in the 1980s provides an excellent overview of these ideas. See also a summary at http://www.infed.org/thinkers/et-schon.htm.
8. For more on Pixar's animation process see http://www.pixar.com/howwedoit/index.html, particularly slide #7.
9. For more on participant structures, particularly as they relate to professional practica, see Shaffer (2005c).
10. Other representations of game play are of course possible. Aldrich, for instance, provides several possible models in his blog-to-become-book *Style Guide for Serious Games and Simulations* (http://clarkaldrich.blogspot.com/. [AQ8]
11. For a short overview of some key concepts from SNA, see http://www.orgnet.com/sna.html.
12. For a more extensive discussion of SNA from an enthusiastic proponent, see Barabási. [AQ9]
13. A key idea in assessment theory is, of course, that the data is not equivalent to the target knowledge. A given player at a given point in time might be thought of as having an epistemic frame that we would like to characterize in terms of the elements of *P*. This is the target of assessment. We can not see this directly, but rather see the things that a person says or does in the situations to provide evidence about the player's epistemic frame—in the same way that we do not see social relationships directly but infer them from behavior at a party.
14. While the data we discuss here comes for coded discourse, we want to emphasize that in principle any recorded data about players in the game can be used with this technique. We, in fact, also use other forms of data in our work, but do not discuss its inclusion in this introductory presentation of ENA for conceptual and rhetorical clarity. Such issues will, of course, be covered in more depth in future publications. [AQ10]
15. Kamada-Kawai is a force-based algorithm for determining the layout of a network graph. The algorithm represents the association between any two nodes as a force—metaphorically the nodes are steel rings and links are springs that connect them. The attractive force of the springs is proportional to the strength of the association, and the algorithm positions the nodes in the layout so as to minimize the potential energy of the springs in the model. Thus, more closely associated nodes appear closer together in the network graph. Only the relative positions of the nodes is significant, not their absolute horizontal or vertical position in the graph.
16. The definition and segmentation of strips of activity is, of course, a significant methodological issue, although beyond the scope of the introductory discussion here. This is an issue that comes up in any analysis of

discourse where one has to specify what constitutes an utterance; here it is of particular importance because the presence of items in the same strip determines the strength of their relationship within the network.

17. These and other network graphs in the paper were produced with the tool SONIA: http://www.stanford.edu/group/sonia/.
18. In the examples that follow, we describe and use graphs that are cumulative from the start of a game. However, it is also possible to use the same methods to create graphs that are cumulative within selected episodes, or tracked with a "moving window" over time. Such graphs might be used to compare frame development in different settings or situations within (or outside) the game.
19. This is partly a statistical artifact—in the sense that early in the game players have not had an opportunity to use frame elements as much as they have had later in the game. On the other hand, experts often act without explicitly articulating the steps of their work (Anderson 1980, 1993; Dreyfus and Dreyfus 1986). Later in the paper we look at metrics (in particular, the concept of *relative centrality*) that address these challenges by quantifying the structure of the network rather than just its overall density.
20. A video of the sequence of frame development is available at http://www.youtube.com/watch?v=pctE4uXimFw.
21. Video of these play histories is available at http://www.youtube.com/watch?v=iUMad85Mulw and http://www.youtube.com/watch?v=goDWYhTVuAE for mentors and players respectively.
22. Task models concern the context and evolving features of the game. In standard forms of educational testing, tasks are discrete and self contained. In interactive assessments such as games and simulation tasks, activity is more continuous, and situations evolve in part in response to actions that the participants take along the way. Of particular importance in creating task models to assess epistemic networks are (a) global considerations such as the epistemic frame that is both the target of learning and the basis of assessment, (b) interface design such that the knowledge components of the epistemic frame are evoked and affordances are provided for the players to interact with, and in fact create, situations that revolve around this knowledge, (c) episodic features such as goals, (d) features of situations that designers can control, and (e) features of situations that must be recognized when they occur, in order to properly interpret players' actions. [0]

References

Adams, P. C. 1998. Teaching and learning with SimCity 2000. *Journal of Geography* 97 (2): 47–55.

Anderson, J. R. 1980. *Cognitive psychology and its implications*. San Francisco: Freeman.

Anderson, J. R. 1993. *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Antráas, P., L. Garicano, E. Rossi-Hansberg, and National Bureau of Economic Research. 2005. Offshoring in a knowledge economy. http://papers.nber.org/papers/W11094. [AQ11]

Barab, S. A., K. E. Hay, M. G. Barnett, and K. Squire. 2001. Constructing virtual worlds: Tracing the historical development of learner practices/understandings. *Cognition and Instruction* 19 (1): 47–94.

Barabási, A. Linked: How everything is connected to everything else and what it means. [AQ12]

Beckett, K. L., and D. W. Shaffer. 2004. *We built this city: Developing students' understanding of ecology through the professional practice of urban planning*. Paper presented at the International Conference of the Learning Sciences (ICLS), Santa Monica, CA.

Beckett, K. L., and D. W. Shaffer. 2005. Augmented by reality: The pedagogical praxis of urban planning as a pathway to ecological thinking. *Journal of Educational Computing Research* 33 (1): 31–52.

Bejar, I. I., and H. I. Braun. 1999. *Architectural simulations: From research to implementation* (No. ETS RM-99-02). Princeton, NJ: Educational Testing Service.

Blunden, B. 2004. *Offshoring IT: The good, the bad, and the ugly*. Berkeley, CA: Apress.

Bogost, I. 2005. Procedural literacy: Problem solving with programming, systems, and play. *Journal of Media Literacy* 52 (1–2): 32–36.

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

Bruner, J. S. 1996. *The culture of education*. Cambridge, MA: Harvard University Press.

Burgess, J., and J. Connell. 2006. *Developments in the call centre industry: Analysis, changes, and challenges*. New York: Routledge.

Carley, K., and M. Palmquist. 1992. Extracting, representing and analyzing mental models. *Social Forces* 70: 601–636.

Clauser, B. E., M. J. Margolis, S. G. Clyman, and L. P. Ross. 1997. Development of automated scoring algorithms for complex performance assessments: A comparison of two approaches. *Journal of Educational Measurement* 34: 141–161.

Csikszentmihalyi, M. 1996. *Creativity: Flow and the psychology of discovery and invention*, 1st ed. New York: HarperCollins.

Delandshere, D. 2002. Assessment as inquiry. *Teachers College Record* 104 (7): 1461–1484.

diSessa, A. A. 2000. *Changing minds: Computers, learning, and literacy*. Cambridge, MA: MIT Press.

Dreyfus, H. L., and S. E. Dreyfus. 1986. *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: Free Press.

Friedman, T. 2005. *The world is flat: A brief history of the twenty-first century*. New York: Farrar, Straus and Giroux.

Gee, J. P. 2003. *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Gee, J. P. 2005. What will a state of the art video game look like? *Innovate* 1 (6). http://www.innovateonline.info/index.php?view=article&id=80). [AQ13]

Gee, J. P., G. A. Hull, and C. Lankshear. 1996. *The new work order: Behind the language of the new capitalism*. St. Leonards, NSW: Allen & Unwin.

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

Goldenberg, E. P., and A. A. Cuoco. 1998. What is dynamic geometry? In *Designing learning environments for developing understanding of geometry and space*, edited by R. Lehrer and D. Chazan, 351–368. Mahwah, NJ: Lawrence Erlbaum.

Hagel, J., and J. S. Brown. 2005. *The only sustainable edge: Why business strategy depends on productive friction and dynamic specialization*. Boston: Harvard Business School Press.

Harel, I., and S. Papert, S, eds. 1991. *Constructionism*. Norwood, NJ: Ablex.

Hatfield, D., and D. W. Shaffer, D. W. 2005. Press play: Designing an epistemic game engine for journalism. Under review by *International Conference of the Learning Sciences (ICLS)*.

Hatfield, D., and D. W. Shaffer. 2006. *Press play: Designing an epistemic game engine for journalism*. Paper presented at the International Conference of the Learning Sciences (ICLS), Bloomington, IN.

Hickey, D. T., S. J. Zuiker, G. Taasoobshirazi, N. J. Schafer, and M. A. Michael. 2006. Balancing varied assessment functions to attain systemic validity: Three is the magic number. *Studies in Educational Evaluation* 32: 180–201.

[AQ14] Hoyles, C., R. Noss, and R. Adamson. 2002. Rethinking the microworld idea. *Journal of Educational Computing Research* 27 (1 and 2): 29–53.

Hunter, I. 2006. *The Indian offshore advantage: How offshoring is changing the face of HR*. Aldershot, Hants, UK and Burlington, VT: Ashgate Pub.

Kafai, Y. 1996. Learning design by making games: Children's development of design strategies in the creation of a complex computational artifact. In *Constructionism in practice: Designing, thinking and learning in a digital world*, edited by Y. Kafai and M. Resnick, 71–96. Mahwah, NJ: Lawrence Erlbaum Associates.

Kamada, T., and S. Kawai. 1989. An algorithm for drawing general undirected graphs. *Information Processing Letters* 31: 7–15.

Kanter, R. M. 2001. Evolve! Succeeding in the digital culture
[AQ15] of tomorrow. Sound recording. St. Paul, MN: HighBridge. Published by arrangement with Harvard Business School Press.

Kaput, J. J. 1992. Technology and mathematics education. In *Handbook of research on mathematics teaching and learning*, edited by D. A. Grouws, 515–556. New York: Maxwell Macmillan International.

Kehal, H. S., and V. P. Singh. 2006. *Outsourcing and offshoring in the 21st century: A socio-economic perspective*. Hershey, PA: Idea Group.

[AQ16] Kirschner, P. A., J. Sweller, and R. E. Clark. 2006. Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist* 41 (2): 75–86.

Koschützki, D., K. A. Lehmann, L. Peeters, S. Richter, D. Tenfelde-Podehl, and O. Zlotowski. 2005. Centrality indices. In *Network analysis: Methodological foundations*, edited by U. Brandes and T. Erlebach, 16–61. Berlin, Heidelberg: Springer-Verlag.

Markusen, J. R. 2005. Modeling the offshoring of white-collar services from comparative advantage to the new theories of trade and FDI. http://papers.nber.org/papers/w11827. [AQ17]

Messick, S. 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher* 23 (2): 13–23.

Mislevy, R. J. 1996. *Evidence and inference in educational assessment* (CSE Technical Report 414). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Mislevy, R. J. 2006. *Issues of structure and issues of scale in assessment from a situative/sociocultural perspective* (No. 668). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Mislevy, R. J., and M. M. Riconscente. 2006. Evidence-centered assessment design: Layers, concepts, and terminology. In *Handbook of Test Development*, edited by S. Downing and T. Haladyna, 61–90. Mahwah, NJ: Erlbaum.

Mislevy, R. J., and L. Steinberg. 2003. *A four-process architecture for assessment delivery, with connections to assessment design*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Mislevy, R. J., L. Steinberg, and R. Almond. 2001. *Making sense of data from complex assessments*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Mislevy, R. J., L. S. Steinberg, and R. G. Almond. 1999. Evidence-centered assessment design. http://www.education.umd.edu/EDMS/mislevy/papers/ECD_overview.html. Educational Testing Service. [AQ18]

Mosteller, F., and J. W. Tukey. 1977. *Data analysis and regression: A second course in statistics*. Reading, MA: Addison-Wesley.

Nash, P., and D. W. Shaffer. 2008. *Player-mentor interactions in an epistemic game: A preliminary analysis*. Paper presented at the International Conference of the Learning Sciences (ICLS), Utrecht, Netherlands.

Noss, R., and C. Hoyles. 1996. *Windows on mathematical meanings: Learning cultures and computers*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Nulty, A., and D. W. Shaffer. 2008. *Digital Zoo: The effects of mentoring on young engineers*. Paper presented at the International Conference of the Learning Sciences (ICLS), Utrecht, Netherlands.

Papert, S. 1980. *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.

Resnick, M. 1994. *Turtles, termites, and traffic jams: Explorations in massively parallel microworlds*. Cambridge, MA: MIT Press.

Rohde, M., and D. W. Shaffer. 2004. Us, ourselves, and we: Thoughts about social (self-) categorization. *Association for Computing Machinery (ACM) SigGROUP Bulletin* 24 (3): 19–24.

Schon, D. A. 1983. *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Schon, D. A. 1985. *The design studio: An exploration of its traditions and potentials*. London: RIBA Publications.

Schon, D. A. 1987. *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. San Francisco: Jossey-Bass.

Serra, M. 1997. *Discovering geometry: An inductive approach*, 2nd ed. Berkeley, CA: Key Curriculum Press.

Shaffer, D. W. 1997a. Escher's world: Learning symmetry through mathematics and art. *Symmetry: Culture and Science* 8 (3–4): 369–393.

Shaffer, D. W. 1997b. Learning mathematics through design: The anatomy of Escher's World. *Journal of Mathematical Behavior* 16 (2): 95–112.

Shaffer, D. W. 2002a. The Design Studio: A promising model for learning to collaborate: Thoughts in response to Hall, Star, and Nemirovsky. In *Computer support for collaborative learning 2*, edited by T. Koschmann, R. Hall, and N. Miyake, 223–228. Mahwah, NJ: Lawrence Erlbaum Associates.

Shaffer, D. W. 2002b. Design, collaboration, and computation: The design studio as a model for computer-supported collaboration in mathematics. In *Computer support for collaborative learning 2*, edited by T. Koschmann, R. Hall, and N. Miyake, 197–222. Mahwah, NJ: Lawrence Erlbaum Associates.

Shaffer, D. W. 2004a. *Epistemic frames and islands of expertise: Learning from infusion experiences*. Paper presented at the International Conference of the Learning Sciences (ICLS), Santa Monica, CA.

Shaffer, D. W. 2004b. Multisubculturalism: Computers and the end of progressive education. In submission to *Educational Researcher*.

Shaffer, D. W. 2004c. Pedagogical praxis: The professions as models for post-industrial education. *Teachers College Record* 106 (7): 1401–1421.

Shaffer, D. W. 2004d. When computer-supported collaboration means computer-supported competition: Professional mediation as a model for collaborative learning. *Journal of Interactive Learning Research* 15 (2): 101–115.

Shaffer, D. W. 2005a. Epistemic games. *Innovate* 1 (6). Reprinted in Computer Education (in press).

[AQ19] Shaffer, D. W. 2005b. Epistemic games. *Innovate* 1 (6). http://www.innovateonline.info/index.php?view=article&id=79.

Shaffer, D. W. 2005c. *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 Research. http://www.wcer.wisc.edu/publications/workingPapers/Working_Paper_No_2005_8.pdf

Shaffer, D. W. 2005d. Multisubculturalism: Computers and the end of progressive education. Under review by *Teachers College Record*.

Shaffer, D. W. 2005e. *Video games and the future of learning* (WCER Working Paper No. 2005–4). Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Shaffer, D. W. 2007. *How computer games help children learn*. New York: Palgrave.

Shaffer, D. W. In press-a. Epistemic frames for epistemic games. *Computers and Education*.

Shaffer, D. W. In press-b. *Epistemography and the participant structures of a professional practicum: A story behind the story of Journalism 828* (WCER Working Paper Series). University of Wisconsin-Madison, Wisconsin Center for Educational Research.

Shaffer, D. W. In press-c. Learning in design. In *Foundations for the future: The need for new mathematical understandings and abilities in the 21st century*, edited by R. Lesh, J. Kaput, and E. Hamilton. Hillsdale, NJ: Lawrence Erlbaum Associates.

Shaffer, D. W. In press-d. *Multisubculturalism: Computers and the end of progressive education* (WCER Working Paper Series). University of Wisconsin-Madison, Wisconsin Center for Educational Research; also in submission to Educational Researcher.

Shaffer, D. W. In press-e. Wag the kennel: Games, frames, and the problem of assessment. In *Handbook of Research on Effective Electronic Gaming in Education*, edited by R. Fertig. http://epistemicgames.org/cv/papers/wagthekennel.pdf. [AQ20]

Shaffer, D. W., and J. P. Gee. 2007. Epistemic games as education for innovation. In *Learning through digital technologies*, edited by J. D. M. Underwood and J. Dockrell, 71–82. Leicester, UK: British Journal of Educational Psychology. http://epistemicgames.org/cv/papers/Ep-games-innovation-Shaffer-Gee-BPS-2007.pdf. [AQ21]

Shaffer, D. W., and K. D. Squire. 2005. The pasteurization of education. Under review by *International Conference of the Learning Sciences (ICLS)*.

Shaffer, D. W., and K. D. Squire. 2006. The pasteurization of education. In *Technology and Education: Issues in Administration, Policy, and Application in K12 Schools*, vol. 8, edited by S. Y. Tettegah and R. C. Hunter, 43–55. London: Elsevier.

Shaffer, D. W., K. Squire, R. Halverson, and J. P. Gee. 2005a. Video games and the future of learning. *Phi Delta Kappan* 87 (2): 104–111.

Shaffer, D. W., K. D. Squire, R. Halverson, and J. P. Gee. 2005b. Video games and the future of learning. *Phi Delta Kappan* 87 (2): 104–111. [AQ22]

Shute, V. J., M. Ventura, M. I. Bauer, and D. Zapata-Rivera. In press. Melding the power of serious games and embedded assessment to monitor and foster learning: Flow and grow. In *The social science of serious games: Theories and applications*, edited by U. Ritterfeld, M. J. Cody, and P. Vorderer. Philadelphia, PA: Routledge.

Shute, V. J., and D. Zapata-Rivera. In press. Using an evidence-based approach to assess mental models. In *Understanding models for learning and instruction: Essays in honor of Norbert M. Seel*, edited by D. Ifenthaler, P. Pirnay-Dummer, and J. M. Spector. New York: Springer.

Starr, P. 1994. Seductions of sim: Policy as a simulation game. *The American Prospect* 5 (17): 19–29.

Svarovsky, G. N., and D. W. Shaffer. 2004. *Berta's Tower: Understanding physics through virtual engineering*. Paper presented at the International Conference of the Learning Sciences (ICLS), Santa Monica, CA.

Svarovsky, G. N., and D. W. Shaffer. 2006a. *Berta's Tower: Developing conceptual physics understanding one exploratoid at a time*. Paper presented at the Paper presented at the International Conference of the Learning Sciences (ICLS), Bloomington, IN.

Svarovsky, G. N., and D. W. Shaffer. 2006b. *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., and D. W. Shaffer. 2006c. *Engineering girls gone wild: Developing an engineering identity in Digital Zoo*. Paper presented at the Poster accepted at the International Conference of the Learning Sciences (ICLS), Bloomington, IN.

Svarovsky, G. N., and D. W. Shaffer. 2007. SodaConstructing knowledge through exploratoids. *Journal of Research in Science Teaching* 44 (1): 133–153.

Wasserman, S., and K. Faust. 1994. *Social networks analysis: Methods and applications*. Cambridge: Cambridge University Press.

AUTHOR QUERIES

TITLE: Epistemic Network Analysis: A Prototype for 21st Century Assessment of Learning

AUTHORS: David Williamson Shaffer, David Hatfield, Gina Navoa Svarovsky, Padraig Nash, Aran Nulty, Elizabeth Bagley, Ken Frank, André A. Rupp, and Robert Mislevy

[AQ1]: AU: Please provide email addresses and titles (where appropriate) for all authors.
[AQ2]: AU: Please supply 3–5 keywords.
[AQ3]: AU: An author name is required here for this 2002 citation.
[AQ4]: AU: Author name is required for this 2006, 2005a citation.
[AQ5]: AU: For all figures, please insert a figure call out in the text naming the figure number.
[AQ6]: AU: I can't follow the foregoing sentence.
[AQ7]: AU: something missing in the phrase "such as item response theory provide measurement errors"?
[AQ8]: AU: Please include access date.
[AQ9]: AU: Please give pub date and complete reference in references list.
[AQ10]: AU: for? from?
[AQ11]: AU: Please give access date.
[AQ12]: AU: Please provide full publication info.
[AQ13]: AU: Please give access date, unless you used hardcopy, in which case please give page numbers above.
[AQ14]: AU: This entry not cited.
[AQ15]: AU: Please add format
[AQ16]: AU: This entry not cited.
[AQ17]: AU: Please give access date.
[AQ18]: AU: Please give access date.
[AQ19]: AU: Please give access date, unless you are referring to hardcopy, in which case please give page numbers.
[AQ20]: AU: Please give access date.
[AQ21]: AU: Please give access date.
[AQ22]: AU: What is the difference between this entry (2005b) and the previous one (2005a)? They have the same names and page numbers??