

MacArthur Documenting Learning Meeting San Diego, February 3-4, 2012

Day 1, February 3

Lunch (12 - 1 pm)

Prior to the events discussed in this video Jay gave a brief summary agenda for the meeting

P4 discusses conversation with Nichole Pinkard about what kind of measures to be looking for in YouMEDIA:

"I have been pushing for measures about the collective use of the space day-to-day rather than individual trajectories of participation and learning." Individual stories are not hard to tell and they can be easily selected, but we are not good at saying "what has the space done at exposing how many kids to how many things and ways . . . it is important to measure . . . to have a site like YouMedia where we have alternative metrics."

Rest of the conversation during lunch revolves around the future of universities in light of the push to have lectures put online. Later the discussion turns to the state of college admissions.

Discussion (1 - 4:30 pm)

Video 2

P4: Given previous efforts at assessment in out-of-school (Computer Clubhouses, 5D) what have we learned so far, what is different about today's new media and educational interventions around it. There is "no sense of best practices or what the palate of practices are for assessing and making arguments about outcomes in more unstructured types of technology settings." Background of project and the question of outcomes

P4: In writing the outcomes section of a proposal she was writing she realized that "we could put anything here." The aim is get what "the state of thinking is" in informal learning assessment. The aim is to develop "a collective statement of what we should be looking for and how we should look for it." This is going to become more important as more and more learning is "migrating to personalized and unstructured settings" and not just settings like after-school and museums.

P4: Since the beginning of this conversation about assessment MacArthur has been shifting its focus from innovation to scale-up (e.g. quest to learn schools and YouMedia are scaling really fast). Many people want to adopt these models. Currently **P4** is working with YouMedia folks to establish what

50 the core design features really are.

51

52 **P4:** The problem: you have people who are good at documentation and
53 assessment, but "unless the design community and [learning theorists] are
54 clear about what these spaces are teaching kids that are different, then you
55 can't have a documentation person come in and do an effective job. So we need
56 to have a better articulation of what the targets are before you can even get
57 measurement people into the conversation . . . [what's missing] is an actual
58 conceptualization of what the targets are.

59

60 **P3:** Targets for the community of interest that is doing something or for the
61 person who is trying to assess that from the outside? A group might have as a
62 goal getting the learner to be good at Photoshop, but the researcher might be
63 interested in understanding how learning how to use Photoshop transfer to
64 something else?

65

66 **P4:** With MacArthur there is the advantage that funders and designers are
67 "kind of" on the same page.

68

69 **P4:** These MacDoc meetings are about consensus by the group of what the
70 outcomes should be.

71

72 **P4:** On the example of Starcraft research ". . . we should have alignment
73 about why we think Starcraft communities are interesting from a learning
74 perspective and why Quest to Learn schools are interesting from a learning
75 perspective. We should have alignment on the values and outcomes that we're
76 looking for, so at that layer . . . we need to stay coordinated although the
77 actual studies are not being run by the same people . . . Connie supports and
78 funds what is going on at MacArthur sites, P4 is responsible for "managing
79 the conceptual framing of how the documentation works . . . trying to see if
80 there is any alignment" without success.

81

82 **P3:** Discusses his perspective on the history of this MacArthur concern with
83 assessment. MacArthur was interested in supporting "a particular kind of
84 learning which we called situated/embodied sociocultural learning . . .
85 people had argued that [features of this kind of learning] were more present
86 in some out-of-school spaces than they were in school . . . The push was to
87 study this kind of learning, but the problem was that it could not be
88 assessed by using the standard 19th century assessment methods of
89 standardized tests. Another problem: Assessment was thought about too late in
90 the process - the need to worry about assessment first in order to put into
91 place activities that can be assessed for what one wants to assess . . . so
92 how do you marry these standardized forms of assessment to this type of
93 learning "so that you revolutionize both."

94

95 **P3:** Macarthur wanted a paradigm change in schools vs. Gates F. wants to
96 quick, big impact change for equity reasons. To get paradigm change MacArthur
97 wanted to fund proof of concept work to put pressure on the schools. It was
98 never the idea that they would go to scale immediately "The idea was . . .
99 not to say that we would bring the stuff from out of school to in school, but
100 we would change the paradigm of school so that it would be receptive to new
101 stuff. . . our thing was to try to say how was assessment working in these
102 situated - embodied things in out of school, how could these models of
103 assessment work if we do change the paradigm of schooling."

104
105 **P3:** What has changed is that some people have "bought into the idea that
106 there is this form of out-of-school learning that involves situated problem
107 solving in communities of some sort . . ."
108
109 **P3:** germane issue to the project: "The whole issue of assessment has recently
110 been married to data mining, adaptive technologies, multiple representations,
111 customization, trajectories of growth by communities that didn't get it from
112 games or communities of learning but are now transected to it . . . they're
113 aware that these "affinity spaces" indigenously collect a lot of data and
114 represent it and give it back . . . so people said 'can we find a marrying
115 point there?' "can we cross that divide into schools" . . . is assessment
116 going to become about manipulation of information back to the person learning
117 . . ." Assessment in this case is integrated into the activity.
118
119 **P6:** Discusses former student Phil Piety who worked in data information
120 systems, saw all the data that exists in the educational system already to
121 look at the processes by which it is mined and used for the improvement of
122 the system, finding that it doesn't work as one would like it.
123
124 **P3:** Data mining as double edge sword: great resource for finding trajectories
125 of growth and great for surveillance.
126
127 **P6:** Returns to **P4**'s expectations: A paradigm change in the outcomes "in what
128 counts as the value produced by these systems because this is where the
129 feedback loop closes in on itself. You gather the data about what is going on
130 in order to move the system in a particular direction . . . what are the
131 criteria for what is a good direction to move in?
132
133 **P6** begins to give background on what took place in prior meetings
134
135 First Meeting - defining scope of what was to be reviewed.
136 Out-of-school semi structured settings - there is some design, notion of what
137 good outcomes would be, some structure for the activity to achieve those
138 goals, more free-choice, play.
139
140 Focus on media rich environments with lots of social interaction involved.
141
142 Dropped focus on everyday life activities and solo activities. Also sports
143 and girl scouts.
144
145 Unit of analysis: learning activity systems over time in their institutional
146 contexts. Activity system: many different kinds of activities going on but
147 they have some kind of designed or structured relationship to each other that
148 changes over time.
149
150 Defining learning. Must include system level and individual learning.
151 Learning outcomes: both cognitive academic and affective, social-emotional
152 learning/development included.
153
154 A NEW LIST OF OUTCOMES? We need outcomes for improving the system alongside
155 individual learning outcomes. Improvements in learning academic-cognitive
156 should stand alongside indices of affective, social-emotional, identity.
157
158 **P3:** Affinity spaces have their own situated understandings (criteria) of what
159 the correct affective social emotional states are (e.g. some are tough love,

160 some are not). They have different attitudes about what it means to interact,
161 what identity is.

162
163 **P6:** Connects **P3**'s comments to the issue of "whose goals." Notes that the
164 current paradigm sees all but the cognitive-academic is icing on the cake. To
165 change the paradigm we need to see all of these things as equal: cognitive
166 academic, socio-emotional, identity at the individual and community levels.
167 If you are going to have a successful system you have to show success on all
168 or most of these.

169
170 **P4:** When you look at the informal, the diverse "then you actually have to
171 make the stronger claim that the affective, cultural, identity stuff is more
172 fundamental than the content specific stuff . . . THE PROBLEM: If your target
173 is content expertise, but the model is interest-driven, youth driven, then
174 there is no way you can actually measure all of the content and knowledge
175 outcomes. So you what you actually have to say is what we are measuring are
176 the fundamental drivers of learning that can be applied to different content
177 domains like resilience or effective learning choices . . . these are
178 dispositions that kids can apply in any domain, but if you chase every piece
179 of knowledge there is no way we will win this fight.

180
181 **P6:** Lines are blurred among these outcomes (e.g. There is content learning
182 that changes people's identities). "A project should identify the emphases
183 and priorities among these outcomes and be judged by how well it achieves
184 those and anything else on the list is icing on the cake." The point is not
185 to make the academic-cognitive the final measure

186
187 **P9:** One thing to look at is to study if one or more outcomes pull the others
188 along. In his own work Identity was the "compass" that lead to the other
189 outcomes. It's not just about solving the problems on the test, it's also
190 about navigating institutions.

191
192 **P4:** It's critical for content focused programs to be measuring learning of
193 the knowledge that the programs is designed to deliver. **The question of**
194 **INTEROPERABILITY: individual programs of course need to be assessed on**
195 **content specific targets, but as a community "we need to develop some shared**
196 **metrics that are not content specific."**

197
198 **P3:** Content is a loaded word: "the stuff that's supposed to be stored in your
199 head or it can be viewed as the problems, or problem skills you need and this
200 gets hidden if we keep using the word content . . . you could call it
201 geometry or you could call it designing in second life."

202
203 **P10:** Let's not dichotomize the cognitive and the affective.

204
205 **P6:** ideally we can say that our notion of learning of outcomes "has to be
206 inclusive of all of these things because you can't really do a good job of
207 defining genuinely valuable outcomes unless you see how they interpenetrate
208 with the other ones"

209
210 **P3:** **But in the out of school spaces the identity issue is not uniform. Each**
211 **space is an "identity niche" . . . what you want to learn is relative to the**
212 **niche.**

213
214 **P3:** Raises the question of open sourcing: many people making little
215 contributions.

216
217 **P6:** Notion of role diversity with different with different balances of
218 criteria depending on what the role is, is important.
219 What is wrong with our standardized system is the standardization -
220 normalization, that everyone should know exactly the same thing about
221 everything vs. having a community where different people with different
222 skills collaborating.
223
224 Defining informal: out-of-school, free-choice, self-motivated, commitment
225 over time.
226
227 **P9:** likes "enjoyment" but not "for it's own sake." Problematic because it
228 implies intrinsic motivation. People enjoy things FOR A REASON (e.g. a
229 community of peers values what you're doing. **P6:** Changes to "ENJOYABLE"
230
231 **P2:** Counterpoint to fidelity of implementation is principled adaptation to
232 local purposes.
233
234 **P6:** Local adaptation notion may link into **P4**'s concern about interoperability
235 and shared criteria and metrics. Features on the slide refer to setting-by-
236 setting, but if we want to start thinking how to take something that's
237 working and adapting it for a new location (scaling up) then you need a new
238 notion of fidelity, one in which the viability of local adaptation is equally
239 as important with the replication of the successful features in the original
240 model.
241
242 **P3:** Two other features that are prevalent out-of-school: good versions of
243 activities are not age graded and time is not as much of an issue (e.g. you
244 don't get credit for being faster - no one cares how long it took you to
245 design something in Second Life)
246
247 **P3:** The issue of interoperability: "Let's say you don't want to standardize
248 the [innovative activities]but I want to say that each one of them formulates
249 certain norms of excellence . . . they make internal judgments about when
250 and where people have met those norms of excellence . . . you can say across
251 different groups that are doing similar things how did the norms compare . .
252 . to other groups or to the standards of the experts."
253
254 **P2:** LIFE center work - trying to synthesize ideas of agency in relation to
255 fields of choices within a setting.
256
257 **P3:** Schwartz - assessing choice. One of the ways to assess where people are
258 in what their doing is by looking at who the person asked for help from, how
259 did they do it, when did they do it. Schwartz showed that choice-based
260 assessment predicts knowledge-based assessment.
261
262
263 **P3:** **Ralf Kostor** "Theory of Fun" Book: if you design something and you try to
264 make an agent out of the player, one thing that the player will do is
265 optimize learning (including resorting to cheating). "Always this tension of
266 'how do I build enough room for optimization' which is a real way for you to
267 show commitment, but not so much that you wreck the whole thing."
268
269 **P6:** This is a design question: how far do you open things up before they fall
270 apart.

271
272 **SLIDE - ISSUES: CRITERIA:** whose criteria; assessment for what purposes; what
273 is valued by participants during the activity; valued by experts, parents,
274 community; consequential for other (including conventional academic learning
275
276 **P3:** notion of participant is problematic - not always clear who is a
277 participant or if everyone is participating in the same thing. E.G. WoW
278 players in the same environments don't play in the same way - what is
279 "community"? You have to describe the trajectory of participants in the
280 activity
281
282 **P4:** About slide, not just "consequential" - participation internal to the
283 community is also consequential.
284
285 **P4:** This is really about criteria for assessment so it's really about
286 external criteria, external to that practice.
287
288 **P6:** Nested circles of criteria - of the participants, of the larger
289 community. **P4:** need for a relational language **P6:** Consequentiality to other
290 learning domains? **P3:** What's another way of navigating - "there's a space of
291 these domains and people navigate differently through them.
292
293 **Day 1, Video 3**
294
295 **P3:** Navigation of a space is valuable not only for where it can take you, but
296 for the navigation itself.
297
298 **P6:** You want to document over a long enough period to say something about the
299 relationship between the amount of time a person spends on an activity and
300 what they go on to do.
301
302 **P2:** Social recognition as a tool for "networking" a learner into the next
303 activity
304
305 **P4:** "Connections to other sites of learning" in place of using the term
306 "consequentiality"
307
308 **P3:** Kingbeach: Transfer as location transfer - mobility
309
310 **P4:** We have many ways of measuring outcomes, what we lack are ways of
311 understanding the mechanism.
312
313 **P3:** Brings up example of how people retrospectively point to moments that
314 they consider critical to their development into a particular career path.
315 The rest of the discussion in this tape focuses on examples of this: Andy
316 diSessa coming to physics through working with his father in the garage;
317 Mitch Resnick designing mini golf courses, **P6** and dinosaurs.
318
319 **P3:** The important thing about if and how people come to identify themselves
320 with activities, careers, fields, etc is that this gives the person the
321 confidence to engage with problems, to feel that they can and should work on
322 a problem.
323
324 **Move to Slides about Meeting 2**
325

326 Day 1, Video 4

327
328 **P6:** Documentation - what's the role and potential necessity of all this rich
329 media. Could you do it without video. Rich media is necessary because you
330 can't anticipate what will happen. You need video for arguing HOW things
331 happened. Field notes may only give you someone's interpretation of how it
332 happened.

333
334 **P3:** Add two questions to this slide: (1) how does the documentation you
335 gather compare to the documentation that is already generated by the activity
336 under study (including information that the people who run the activity
337 generate for themselves for self-assessment); (2) can you really make claims
338 about what happened if you haven't participated for a long time yourself?

339 WoW example: you have to have lived through at least three patches because
340 the patches embody the theories of changes that needed to be made to the
341 world.

342
343 **P6:** We could say what the ideal data set but then would it be practical to
344 actually analyze it (e.g. 200 hrs. of video)

345
346 **P3:** Even if you don't use video you still have to contend with massive
347 amounts of data (e.g. theory crafters in WoW). The issue is not the data but
348 how you represent it.

349
350 **P6:** We need not just theories for longitudinal research but also
351 methodologies. The example of Deb Roy, **P4T**, massive documentation of the
352 birth of a child's word . . . Self-documenting activities & in-progress
353 versions of the products that learners produce . . . Affective as key
354 indicator of the success of the learning activity system (e.g. staying up all
355 night, wanting to keep doing the activity). . . How to argue an evidential
356 chain (change?): how to link the affective evidence to one of your outcome
357 criteria. (Unless you take the affective as learning evidence in its own
358 right)

359
360 **P2:** Are there physiological things that help you see these powerful affective
361 experiences? What is the meaning of the emotion in relation to what one wants
362 to accomplish?

363
364 **P3:** It's not that the games are "fun" it's that the learner persists past
365 failure, it's the engagement, you don't quit and all of this requires
366 affective engagement. Fun is relevant, but the feeling of mastery is also
367 key. The issue of frustration: you need some in order to sustain motivation.

368 **P6:** There is a positive value to both negative and positive emotions.

369 **P3:** But there are criteria for productive vs. unproductive frustration.
370 People have to feel the task is doable, that there is help available, and
371 they have to get feedback about the progress they're making. Failure gives
372 you the feedback.

373
374 **P6:** Other **negative emotions**. Barab video game example: River of Justice.
375 Concerns civil war in Africa and the difficult moral choices one has to make
376 under those circumstances. **P6** played the game and at one point in the game
377 the player is confronted with the choice of shooting an innocent child or
378 being killed. Consider the perspectival shift that results from this
379 experience, how it shapes the way a person interprets events in the rest of
380 the game (e.g. reconciliation phase in the game - meeting the mother of the

381 murdered child) - "there are other forms of learning that open up when you
382 are on the other side of this that would not open up emotionally and
383 intellectually if you don't go through this.

384
385 **P10:** Engagement as a holistic construct. **P3:** Engagement is the marriage of
386 the cognitive and the affective. **P2:** NRC report, strands of learning that
387 intertwine (separate but connected) **P3:** Engagement and motivation as terms
388 that need to be revived, it is a social achievement.

389
390 **P6:** Documentation is partially solved by activities in which the participants
391 create analyzable media. **P3:** Learners make these products with respect to
392 standards that are co-constructed with members of their local group. This
393 connects to the issue of feedback: feedback from your peers (can be negative
394 and tolerated), feedback from others **P6:** SB Heath and different forms of
395 feedback

396
397 **P7:** Lisa Soep - Youth radio project

398
399 **COFFEE BREAK**

400

401 **Day 1, Video 5**

402

403 **Slide: PROPOSALS AND ISSUES.** System level learning - e.g. documenting changes
404 in social networks and distribution of practices across networks.

405

406 **P2:** Bill Penuel has been funded to do this kind of work (W/ Falk and
407 Dierking) - agent based modeling. Presentation to be given the following day.

408 **P6:** Organizational learning is one field that has looked into this.

409

410 **P3:** e.g. SIMS building. Early on in the history of the game there were no
411 good tools for building in the world. One had to learn on her own if one
412 wanted to build. Over time more user friendly tools were developed and this
413 allowed more users to participate, more creativity, but it diminished the
414 kind of status that people would gain from going off on their own to learn
415 and develop building tools. Research question: What happens when the old-
416 timers loose their status when these new more user-friendly tools emerge?

417

418 **P6:** Jan Plass and embedded assessment - "**incremental-challenge design model
419 of assessment**" Designers need to have a good sense of what the learning
420 outcomes are in order to meet goals, but what gets missed in this kind of a
421 model? Outcomes that may not seem relevant for achieving the established
422 goals may be consequential for other aspects of the game.

423

424 **P3:** With these incremental challenge approaches one has to consider how
425 algorithmic the achievement is. If it is algorithmic then it can be built in
426 such a way that accomplishing the task can be read as success. But a good
427 game has to allow for multiple solutions. **Valerie Shute model of Stealth
428 Assessment designed to assess the multiple solutions that players develop in
429 their game play. P9: Using the Shute model for YouSTEM project.**

430

431 **P3:** Players have aesthetic criteria for solving problems. This is something
432 that we need to pay attention to. **P6:** Example in math. Efforts to find
433 elegant solutions gave birth to new methods for solving other problems

434

435 **P4:** Can these proposals be mapped to some of the earlier categories that Jay
436 presented? Jan's work is about "assessments internal to the practice that are

437 **being thrown off by the actual activity, but they don't necessarily help us**
438 **to get at the interoperable pieces, or the transfer, mobility.**
439

440 **P3:** They do at a genre level. In any of these activities or games if you have
441 done the algorithmic assessment for say a real-time strategy game then you
442 have learned more than the game, you've learned the genre. But you loose the
443 interoperability when you cross the genres.
444

445 **P4:** But you'd have to design the interoperable widget between the two games.
446 E.G. if you learn the genre of real-time strategy war games your learning
447 curve for a new game is completely different when you take up that game, but
448 you wouldn't know this from an analysis of the metrics internal to each game.
449 The assessor would need an additional metric. **P3:** Game interfaces are
450 similar, so facilitating the transfer.
451

452 **P4:** You need to have the depth and domain specific metrics, plus the
453 interoperability metrics. The proposal on the slide concerning documentation
454 of systems-level learning via changes in social networks and distribution of
455 practices across networks is one that could be an interoperability metric,
456 not domain specific. E.G. If a kid gets good at a game do you see a
457 systematic shift within their network that say clusters around a genre,
458 technology.
459

460 **P6:** Re-analysis potential. E.G. video allows for re-analysis from different
461 perspectives. "It is a criterion of a value of a documentation mode or of any
462 assessment procedure that it can allow other people to go back to the same
463 data and re-evaluate using different criteria. Some kinds of documentation
464 and assessment, their criterion are so outcome-specific that they're no good
465 for anything else. Others are presumably very good, and some may be so good
466 for so many purposes that they're not that good for any one purpose."
467

468 **P3:** E.G. heat maps in Halo, used for different kinds of analyses by players
469 and designers **P6:** And you can't really predict how people will analyze or use
470 these new tools.
471

472 **P3:** The problem of proprietary interests and privacy issues in assessment.

473 **P6:** We are in favor of rich media, but as soon as we use it then we have to
474 think about all the IRB issues. More general discussion of IRB issues (**P2**,
475 **P9**)
476

477 **BADGES Slide**

478 Episode 1: A New Hope - Badges can give wide recognition for non-school
479 learning
480

481 **P3:** Explains history behind Badges. (1) Eva Baker as AERA president called
482 for a system like this (called Qualifications at the time) (2) As gaming
483 systems improved, games were required to have "achievements" within the games
484 and people were blown away by how motivating these achievements were to
485 getting players to persist past failure. So badges are seen as motivational,
486 and an opportunity to honors skills that libraries don't. **P6:** Adds that the
487 open-source programming community came up with an internal system of this
488 kind. In addition there is the notion of being able to identify and seek help
489 from someone who has expertise in a particular domain.
490

491 Episode 2: The Empire Strikes Back - National standards for endorsing local

492 badges? **P6**: Cites example of someone at the DOE seeing badges as something
493 the DOE: this is a potential backdoor to centralized standardization.
494
495 **P2**: LIFE center is leveraging badges in relation to standards. Science
496 literacy - you do want people to understand climate science. You can use a
497 **performance portfolio system** to achieve this kind of assessment.
498
499 Episode 3: A Phantom Menace? - How to keep community-control of badge systems
500 **P6**: The positive in maintaining community control over the development and
501 dissemination of badges - respecting the autonomy of the community while
502 maintaining transparency. The danger is in the scaling up to a national
503 certification system.
504
505 **P3**: Any new device has the potential for good or evil. Academics have little
506 influence because they seem only to focus on criticizing and deconstructing
507 the tool for it's potential evil uses, and not offer an alternative, hence
508 the gap in alternative assessments. Badges could help destabilize the idea of
509 centralized control by promoting a perspective in which different people are
510 expert at different things (vs. forcing everyone to have the same knowledge)
511
512 **P3**: The problem of access. Wealthy families can by their children access to
513 things that schools do not provide (e.g. computer camp). If there is a
514 national mechanism for badges then this might create opportunities for access
515 to those without it. Without the system, then the low-income kid can't get it
516 anywhere. The real evil is that the out of school time system has been co-
517 opted by the wealthy. It is funded by the rich parents and provides a 21st
518 century curriculum for the 1/5 of jobs that are decent.
519
520 **P7**: But **P3**'s example is about access to resources, not about credentialing.
521 **P3** & **P4**: It's both **P4**: Badging doesn't solve the problem of access to
522 resources. **P3**: It can solve the problem of access in the following way: A
523 library or community center agrees to offer a "hacker badge" then the low-
524 income kid can compete with the wealthy one. Both have access to the badge.
525 Without the badge system then only the very rich are left with access to the
526 credentials that really count.
527
528 **P6**: In this scenario what is important is credible endorsement, but from
529 whom, at what level? What is dangerous is the possibility of a centralized
530 system. **P4**: The open badge system is made so that there isn't one central
531 certifier. The dominant criticism that **P4** has seen has been about the dangers
532 of being able to game the system From reviewing the latest batch of Badge
533 proposals **P4** sees that the dominant frame for the projects is gamification
534 and extrinsic motivation.
535
536 **P3**: Cisco game example. Game to teach students about networking and routing
537 (Cisco Mind Share Learning Game for CCENT™ and CCNAÆ) If student completes
538 this and gets a Cisco badge no one would think that the DOE's badge would be
539 better.
540
541 **P6**: **Are there any reasonable conditions to be put on the system to lessen the**
542 **chances of it being taken over and becoming standardized.** **P2**: Nichole's move
543 in iRemix is to give control of the development of the badges to the
544 learners. The learners "control their resume"
545
546 **P3**: Another danger is the dumbing down of badges. The badge has to have
547 validity. **P6**: We do need a reputation system - crowd system of authority.

548
549 **P3:** The fact that all our evaluative systems can be gamed by the wealthy is
550 no reason to not try out these systems. The worst that could happen is that
551 we get a system that replaces grades (i.e. we can't do worse) **P7:** But we have
552 to be mindful of the danger that the badge system could disproportionately
553 direct resources away from those who need them. **P4:** What **P7's** point raises is
554 that you can't look at the mechanism as purely in the design of the badging
555 system, you have to look at the social architecture around the badge system.
556 The problem with tech centered solutions is that people often forget that
557 other piece. So this is a risk with the badging system, that people might see
558 the system as a silver bullet.

559
560 FINAL SLIDE: Beyond Badges - What forms might the crowd-sourcing of
561 assessment take for diverse learning communities? What methods of assessment
562 are most useful for improving learning activity systems as such?

563
564 **P3:** Death of academic disciplines in the move to crowd sourcing around
565 problems. **P6:** This also connects to the notion of collective intelligence.
566 Levels of achievement related to collective intelligence can never be part of
567 a standardized curriculum or assessment for everyone. **P7:** Collectivities need
568 to know who has the expertise. **P3:** But we still need to recruit the expertise
569 of even the most minor contributors - Crowd sourcing chess example.

570
571 **P6:** This is going to be the cutting edge in the future: how to organize
572 collective intelligence. **P3:** Reinvention of Discovery, book by Michael
573 Nielsen. **P6:** Rainbow's End, science fiction book. **P4:** Creating a Class, Mitch
574 Stevens Book

575
576
577 Day 1, Video 6

578
579 **P4:** Question about the clustering of goals of the assessment: "Right now the
580 outline is really around methods of data collection. It's not driven by a
581 sense of the purpose of the (doesn't finish sentence). The methods key to the
582 kind of program which is also tied to the goals of the assessment, that's why
583 I was pushing on the outcomes . . . if your accountability is to just
584 demonstrating domain-specific knowledge it's a very different task than if
585 you're trying to document mobility of learning . . . I really just want to
586 know what's worked for what kinds of programs. I don't need to know
587 everything in the world that's been done but I need to have some orienting
588 markers of 'look, these kinds of programs were assessed in these ways, these
589 are the limitations, these are the successes in terms of documenting certain
590 kinds of outcomes -"

591
592 ST walks in. Small talk.

593
594 **P6:** For the report, at the end of the intro, we should include points that
595 have been made at these meetings as a way to alert the reader, followed by
596 the review of the literature which in some ways will be structured in the
597 ways that the literature structures itself (methods of assessment and what
598 they've been applied to).

599
600 **P5:** So instead of having orienting categories driven by method, these
601 categories should be defined by outcomes?

602
603 **P4:** "That's what I'm thi- just from a user-friendliness kind of approach. You

604 had at the beginning like individual outcomes, you had a few categories which
605 were very useful I thought in orienting - I mean it was useful for me like in
606 just orienting - 'oh, OK these are the really big buckets, these kinds of
607 things, or you're looking at systemic change, these are the categories of
608 things that you want to find out and then these are the kinds of methods that
609 have been applied. It does feel sometimes 'apples and oranges.' You're diving
610 into a conversation that's not really relevant strategically to where you're
611 at."

612
613 **P2:** Discusses how the assessment chapter in the NRC volume on Learning
614 Science in Informal Learning Environments is structured in terms of six
615 strands of science learning as a potential model.
616

617 **P4:** Perhaps structuring the report as "a narrative flow moving from the more
618 conventional domain-specific assessments and then ending with the sort of
619 social networky kind of crowd stuff that's more experimental that we have no
620 track record of yet. that could be very useful because it's also a spectrum
621 between you're looking for something practical, off the shelf that you can
622 drop into an environment vs. you're a researcher and you're developing new
623 documentation methods, psychologically it's a very different orientation. As
624 you (Jay) were going through your presentation I was thinking 'y'know that's
625 exactly my frame because I have a really practical problem which is like 'Ok
626 so - I am not going to be able to do a social network grab around YouMedia,
627 but I could imagine XYZ.

628
629 **P6:** Some of these things would be easier if we already knew answers like what
630 are the clusters what are the categories of these sorts of learning activity
631 domains, categories that make sense because they're unified by the fact that
632 the same outcomes are produced by them, or the same kinds of assessment
633 approaches are successful in them. That's a kind of research level that's
634 beyond what we can do.
635

636 **P3:** You can't draw definitive conclusions until the research is done.
637 How about using specific cases (e.g. discuss an interest-driven site and
638 explain how the issues are played out in this context)
639

640 **P6:** Part of what we what to do tomorrow is establish what kinds of
641 recommendations to make in the report. Some of these may be recommendations
642 about the need to answer some of these kinds of questions and to have
643 research that can tend to do that. "**P4** you seem to be involved in an
644 enterprise that amounts to program evaluation across different instances" so
645 perhaps there needs to be an actual research component to these to see what
646 we can learn about assessment at the interoperability level. Not just to do
647 it but to figure out the conceptual frame for doing it."
648

649 **P4: Wish list.** "The kinds of things that would be incredibly useful . . .
650 things at the research programmatic level would be helpful: so this is what
651 we don't know, this is what we need to know if we're going to do this
652 effectively. . . So this is the piece that I think you don't want to do but
653 it would be really valuable is more of the 'what's worked' kind of thing. If
654 there are a set of recommendations based on the review in terms of what has
655 worked - and you can imagine what has worked being framed in a lot of
656 different ways. But, I just started doing a lit review of different
657 assessments of interest-driven learning programs - of long term effects of
658 participating in arts programs and robotics programs - they're all over the
659 place. They're apples and oranges. I really would like to have an expert

660 voice to help me orient to that whole range of apples and oranges. . . So
661 there's sort of those kind of orienting marks 'like, these are the outcomes
662 that we really think you should be looking for these programs. And so I've
663 listed up things that came out of our discussion like:

664
665 production
666 resilience
667 persistence
668 effective learning choices (which I call attention management)
669 credibility assessment
670 translation and mobility of learning

671
672 These are the things that I intuit as the kinds of outcomes that are coming
673 out of participation in these activities, but I really don't know who studies
674 them or whether there's any evidence beyond the sort of emic evidence - I
675 read Jim's stuff and I can say "oh yeah, Jim is finding the same thing as me"
676 but it's not systematic.

677
678 **P3:** There's very little systematic research other than ones that have been
679 kind of set up - A lot of robotics programs for example. But in the more
680 indigenous ones there are so many issues about how to evaluate and what
681 they're doing that it's just the beginning.

682
683 **P4:** So even like an orientation to say 'Look we have a bunch of qualitative
684 work that is suggestive of these outcomes but what we need is a program to do
685 the interoperability or the testing - like what is missing? Why are we
686 failing so miserably at making a sustained case about the value of these
687 programs, like what are the missing pieces."

688
689 **P7:** We're coming at this from the attitude that we have a treatment and then
690 an outcome and I don't think that applies. Part of the problem here is that
691 we need to think about the nature of the experience/environment and what's
692 happening in that context and that's very hard to do . . . we need to talk
693 about affordances of learning environments and that's language the world is
694 not ready to really hear. So how do we translate that into language that
695 works to make the case for giving kids these kinds of experiences?

696
697 **P3:** Some of these communities produce things that overlap with what some
698 profession produces. E.G> Fan Fiction or Robotics communities. So you can ask
699 are these groups giving rise to high standards? **P10:** We need to look at cases

700 **P4:** There must be some historic examples like 5D and Computer Clubhouses that
701 have been looked at from so many different angles. This is the kind of thing
702 that would just be useful. **P7:** We have to figure out what the little bit that
703 we know amounts to.

704
705 **P4:** "I think recommendations would cluster around like #1 based on our
706 imperfect knowledge today, how do we asses these programs, what are the best
707 tools out there for us to do a practical, accurate documentation, what should
708 we be capturing, what are the opportunities, and then the second is the more,
709 the bigger research questions and gaps if we really want a much bigger
710 paradigm shift in how we do this stuff. And I think both of these could be
711 hugely useful, because even the first one is overwhelming for anyone who is
712 facing this problem of a - what the hell to do.

713
714 **P9:** re: the wide and non-relatable ways that these things have been assessed.
715 Instead of trying to find some uniformity and best practices, instead invite
716 people around the table to go in and try to answer that question as a

717 research project, who start with this conversation as common ground and come
718 together and try to form that across diverse cases.
719

720 **P6:** The fact is that there's just a lot of things here we just don't know. It
721 is important to identify and specify those, talk about some promising
722 directions, but also to see what are the things we do know. There are things
723 we agree on based on a lot of thoughtful experience, but how convincing would
724 that be to people who don't want to believe us.
725

726 **P4:** So you had a question of audience. I don't think this is for the
727 skeptics. I think the people who are already doing this kind of work need
728 (help). We are not at the point of convincing skeptics. I wouldn't even try.
729 I think it's more - if I asked you based on all the intakes you've had, what
730 are the three things you would do to assess learning in new media, what would
731 they be. I need that now and this community has been debating and hashing
732 over this stuff for decades and we keep coming to this point where 'it's so
733 diverse, it's heterogeneous, and blah blah blah" and in the meantime these
734 programs are having shitty assessments done of them, and that's my
735 frustration. I look at what (???) is coming up with and they're trying really
736 hard but it's not with even a fraction of what we're talking about here.
737

738 **P6:** Compares this to the group that produced the white paper on video
739 research in the learning sciences. NSF asked for best practices, but we ended
740 up giving them guidelines for good practices, but then we did a kind of
741 expert consensus about the sorts of things you should be paying attention to
742 and what you should consider doing."

743

744 **P2:** If this is for people who are trying to figure out what to do in the
745 context of their programs, use worked examples of somebody finding their way
746 through that process within families of different kinds of work that are in
747 the portfolio, Maybe not part of the report but something that comes off of
748 it?
749

750 **P9:** You called them shitty assessments, but they come out of shitty
751 assessment producing practices developed by people who don't share theories
752 of learning like the people in this room do. Think about museums. They use
753 surveys to get satisfaction data so they can bring this to their donors. "If
754 we trace things back through the social practices of what evaluation
755 assessment are used for it makes sense for me to scroll forward from people
756 who share certain commitments to theories of learning, and then have them go
757 into particular settings (trails off). Assessment has historically been
758 treated as beneath doing the basic research on learning.
759

760 **P4:** I think that there is openness to trying other things, so throwing in a
761 few tools from this community could be helpful. **P6:** The bar is gradually
762 rising, the standards to demonstrate accountability are rising. But where is
763 the textbook for doing this kind of assessment.
764

765 STOP FOR THE DAY.
766

766 **Day 2, February 4**

767

768 **Discussion 9:00 - 1:00 pm**

769

770

771 **P6:** Provides summary meeting events from the prior day, and agenda for the
772 day. **Audience of the report:** The community of people, researchers and
773 practitioners, who work in informal learning settings. These settings include
774 after-school programs, community centers, and museums. No solo activity, no
775 learning in everyday activities (e.g. home) **Aim:** Offer researchers and
776 practitioners how to perform documentation and assessment of learning in
777 their programs in ways that improve upon existing methods. **Defining learning**
778 not just of individuals, but of the project, of the community, and learning
779 not just as cognitive-academic, but also socio-emotional development,
780 identity development. **Today's agenda:** Focus on the report and case study
781 descriptions from **P8** and **P2**.

782

783 Report structure: Should it be organized around methods, or around outcomes?

784

785 **P1:** Notes **P7**'s work with Bronwyn Bevan on looking at across programs in many
786 sites that are trying to do similar things. How could this work be included
787 as part of the report.

788

789 **P6:** What kinds of recommendations about good practices for documentation and
790 assessment can apply to different sites, allow for comparability to different
791 sites, but with the understanding that they have to be locally adapted to
792 those sites.

793

794 **P5** PRESENTATION OF THE OUTLINE

795

796 **P5** projects detailed outline of the report

797

798 **P1:** Bridig Barron's work of technobiographies in Computer Clubhouses -
799 reconstructing student histories through their design artifacts. This
800 connects to Carol Lee's proposal that we need a better accounting of diverse
801 developmental pathways.

802

803 **P1:** Learned from Nichole Pinkard that in Chicago some after-school programs
804 are sharing data in order to keep track of student who move across sites. **P8:**
805 Key thing to explore is not just learner engagements over time, but the depth
806 of these engagements **P6:** Important general question: We generally look at
807 learner participation within the framework of the activities interior to the
808 project, but this does not tell the full developmental story. Learners are
809 moving in many places in the community. So you need data sharing in order to
810 capture what is happening with the learner in and out of the activity under
811 study.

812

813 **P8:** Milbrey McLaughlin's Youth Data Archive: Has negotiated data agreements
814 with different social services. In exchange for the data sharing they get to
815 define the questions that the data permits asking. Milbrey's group does the
816 analysis. Not Fine-grained, but interesting gross patterns. **P6:** This appears
817 to be a difficult method but rewarding - should be part of the repertoire
818 included in the report. Case studies that reach outside of the project - in
819 the home, school, the peer group. **P6:** One of the outcomes that was defined
820 yesterday concerned the building of connections between what people are

821 learning in the project and what they are learning anywhere else in the
822 universe of their lives.

823

824 **P1:** Example of following learners engaging in STEM learning across many
825 settings (ethnography). Observing that one particular learner deconstructs
826 all the settings (architecturally) and the connection of this to this
827 learners participation in robotics and other science activities. **P1**
828 highlights the fact that this learners tendency to deconstruct might be
829 interpreted as bad behavior depending on the context he is in (e.g. non-
830 science, but classroom contexts). We need to orient to the broader texture of
831 what kids are trying to accomplish. **P6:** Not just extending across settings
832 but across time.

833

834 RETURN TO OUTLINE

835

836 **P1:** Digital story telling as standardized. You can download manuals, etc.:
837 <http://www.storycenter.org/>

838

839 **P6:** We should consider highlighting in the report what is new and
840 experimental, what has more use, is more established. Also, to what extent is
841 the method very specific to what is going on in this project or is it a
842 method that could be used in several different projects.

843

844 **P3:** What **constitutes evidence**? This has been a growing concern at MacArthur.
845 What will make the claims seem valid to a wider audience than us.
846 There is a tension here: Many of the methods we are discussing here are
847 contextually specific methods but they don't speak to validity in
848 psychometric terms. **P3's** solution: **evidence centered design**. This is a
849 tradition is at least accepted by psychometricians and Mislevy (who helped
850 develop the approach) has applied to situated, fine-grained analyses. Valerie
851 Shute builds ECD models for Quest to Learn schools, Quest Atlantis. . . E.G.
852 with the Hull Digital Storytelling example of digital story telling the
853 public will ask what the evidence is that this has had an impact that this
854 person has learned something. ECD is one way to resolve this tension. It is a
855 theory that says worry about collecting evidence for you design at the
856 beginning rather than at the end. This helps you address the validity thing.
857 We need to address this **tension over validity and evidence** in the report in
858 order to have a bigger audience

859

860 **P8:** Traditional methods that resonate with this work would be longitudinal
861 research. **P3:** Jere Confrey Math Learning Trajectories. You do have to tell
862 MacArthur to do longitudinal work, but they need to combine rich interpretive
863 data and microgenetic analyses. The things we study are activities that
864 already produce their own evidence that allows researchers to come in without
865 doing much disturbing of the local ecology.

866

867 **P3:** See Valerie Shute's work where she has gone and actually used the tool of
868 ECD to break down systems thinking (a 21st century skill) to provide a valid
869 working, analyzable definition.

870

871 **P5** continues discussing example from the outline. Reaches Constance
872 Steinkuehler's work **P3:** Steinkuehler research showing that boys who play WoW
873 and do poorly in school are reading 5 grade levels above what they are
874 expected to if given books on WoW.

875

876 **P8:** Angela Barton work connects to community-based programs in STEM. **P1:**
877 **Megan Bang uses** DBR focused on governance on the community side so that the
878 people in the community build what they actually need.
879
880 **P3:** Evaluate learning by evaluating talk - have you picked up the specialist
881 languages. **P6:** When you have a specialized discourse you can look at this
882 across contexts, but if you develop a coding scheme to look at your data,
883 that is not going to transfer. **P3:** If we know that some feature correlates
884 with learning, then you can go in and show that you get this in a particular
885 settings. That is use what the literature says about feature-learning
886 correlates as markers. **P8:** Kathy O'Connor & Sarah Michaels "Accountable Talk"
887 - standardized tool
888
889 **P3:** In game-based research often the realm where people are articulating
890 their understandings are not in the games but in the interest driven sites
891 like online chat forums.
892
893 **P5** discusses the messiness of organizing the report by methods, and opens the
894 discussion to the possibility of organizing the report by outcomes or, as in
895 the NRC report, strands. **P8:** If you could organize it in terms of claims that
896 you want to make about the environments.
897
898 **P6: The tentative scheme that emerged the day prior: Individual learning**
899 **outcomes vs. Collective improvements of the activity system itself**
900 **Then for each: cognitive-curricular-content; social-emotional-identity**
901 **development outcomes. P3:** The social aspect of all these activities creates
902 mid-levels, e.g. in WoW individual, team, and community learning **P7:** Seems
903 unavoidable that we'd get into the "old knowledge-skills-attitude kind of
904 breakdown" if we want to be outcome-centric rather than claims-centric.
905 **P8: Why not learning in social practices? P7: Agrees that it should be**
906 **practice-centric P6: We need to keep the number of these categories relatively**
907 **small. So what if: Individual, intermediate, project or community level;**
908 **Cognitive-curricular-content; social-emotional-identity development outcomes**
909 **Then saying in each of the boxes in this 3 X 2 table which methods might be**
910 **promising.** Then the descriptions of the different examples would site their
911 applicability in the different settings.
912
913 **P3:** It's important to rehabilitate the term knowledge. Use it and stress
914 knowledge as a verb. **P6:** Use "know-how" and talk about knowledge as
915 procedural and declarative knowledge as a special case of procedural
916 knowledge. **P1:** Content obsession of science education remains. Using
917 knowledge in practice perhaps. **P3:** Emphasis has to be placed on equations and
918 content as tools in order to maintain the processual and practical sense of
919 learning in science: you have to be able to apply the equation, not simply
920 regurgitate it.
921
922 **P6:** The issue of audience. We are not out to persuade psychometricians and
923 the content people. We are out to provide assistance to the members of our
924 own community who need guidance on how to do better documentation and
925 assessment of these kinds of things that they do.
926
927 **P8:** Reminds us of **P7's** point that these folks have to be responsible to their
928 funders and to the more traditional folks. **P6:** ECD provides a good
929 recommendation to counter this, plus we will provide a summary of this
930 meeting where we will include this discussion about what people consider to
931 be valuable knowledge (and that in our case we focus on processual knowledge,
932 doing)

933
934 Continue review of outline.
935
936 **P7: Thinks the report as organized by method will not be useful unless we**
937 **attach it to some orienting device. P8: This should also be attached to**
938 **people - who do you go to, who can you email to get help with this. P7: It's**
939 important for smaller institutions to be able to connect to the work of
940 larger institutions.
941
942 **P3:** We should concentrate on claims because what we're making claims about
943 isn't all that clear. We don't have a nice description of the natural
944 categories of what we're doing, these learning systems. This needs to be
945 mapped out.
946
947 **P3: P4** is interested in learning how people learn in their natural
948 environments vs., the more designed environments, ones designed to "get an
949 effect." For **P3** this distinction was never really true, or at least it is
950 less true now: All spaces are designed. As you want to evaluate learning
951 claims to what extent are you going to intervene in what they've already
952 done? E.g. modding. Wouldn't you want to know if modding is in the zpd for
953 some people? Then you have to intervene (interventional ethnography). Why not
954 build in ZPD tests in the game?
955
956 **P1: DBIR as a method - a recommendation**
957
958 **BREAK**
959
960 **Day 2, Video 2**
961
962 **P2 presentation**
963
964 **P2:** Background doing robotics and physical computing workshops with kids
965 (Northwester, **P4T**, Stanford). Was bothered by the fact that in these initial
966 workshops, which lasted 1 - 2 weeks, the process was lost (all that was left
967 were pre/post interview data and artifacts). Also bothered by the attitude by
968 some that it was sufficient evidence of learning that a child had produced an
969 artifact like a robot or a program, but people didn't pay attention to the
970 actual detail of the work. All of this motivated **P2's** move to **learning**
971 **analytics** and instrumenting spaces (fab lab) to capture the process of
972 production. Kinds of instrumentation: panoramic cameras, skin conductivity,
973 microphone arrays, mini gps arrays, experience sampling.
974
975 **P1:** Project (progress?) portfolio tool, Chris Quintana mobile tool (Zydeco)
976
977 **P2: Multimodal learning analytics:** More discussion of techniques: Screen
978 capture; computer vision techniques to detect faces to detect where in the
979 space participants are located and to perform object detection to study how
980 people manipulate and experiment with objects (e.g. seeing how people engage
981 in trial and error at first, then are suddenly able to put objects together
982 at a later time.) Combining instrumenting techniques and machine learning is
983 motivated by the hypothesis that there are underlying structures in how
984 people experiment with objects (pattern recognition). This work is informed
985 by interviews with people at different levels of expertise (engineers, phd
986 students, undergrads). The interviews also include design tests. This data
987 is used to assess if there are markers of engineering expertise that go
988 beyond knowing equations, etc. (i.e. learning by doing).
989

990 The interview & design test data is used to compare with what the kids do in
991 the instrumented spaces. A key example: Redundancy is something that tends to
992 show up primarily in expert work. Error analysis. Interviews with kids are
993 studied to establish when the kids begin to incorporate things like
994 redundancy, error analysis. Example from analyses of discourse: Relative
995 comfort/discomfort with silences - experts comfortable with silences, novices
996 fill the silences with talk. Context: month-long work with kids during
997 intersession, 6 hours a day.
998
999 **P6:** Asks about the details of the interview component. **P2:** Pre/post and in
1000 between interviews of different formats (self-audio record, researchers
1001 interviewing kids as they work, & in-room interviews not in work (design
1002 tasks during the interview)). Example question: what's the first component
1003 that would fail after using the energy saving device that you created 100
1004 times?
1005
1006 **P8:** What is the process of moving from the raw data to making an argument
1007 about how this relates to learning? **P2:** Sometimes the focus is on the purely
1008 quantitative data (e.g. experience sampling and survey data) **P8:** Asks about
1009 the convergence of the biosensing data with the other data. **P2:** Small N for
1010 biosensing data because of participant concerns about the dangers of the
1011 instruments. As a result, **P2's** team performs case studies on the few
1012 participants who agree to provide this kind of data. Biosensing data
1013 collection requires interventions in activity for calibration.
1014
1015 **P2:** Interested in issue of motivation. What are the optimum levels of
1016 frustration? Discusses findings suggesting that having a neutral experience
1017 at the moment when you begin working on a task is less likely to lead to
1018 learning than if you have a positive or negative experience.
1019
1020 **P6:** **P2's** is a case of ideal documentation. Re: outcomes, the model here seems
1021 to be one of comparing the outcomes to those established for experts.
1022
1023 **P6:** What about unexpected things - do you see learning that was unexpected?
1024 **P2:** One surprise related to behavior throughout the activity. "Misbehaving"
1025 students appear not to be accomplishing anything, but when interviewed they
1026 demonstrated that they could "talk and work at the same time." . . . Discusses
1027 an example of the importance for the kids of connecting what they do in the
1028 activity to things in their daily life. Nametag example - creating the
1029 nametag for no real reason vs. doing it for someone they cared about. Latino
1030 vs. Anglo split: Latino's demonstrated more dedication to the task when they
1031 connected the task to their family (e.g. making the tag for a parent) vs. the
1032 Anglo kids seeing the "no real reason" task as "something for me" and
1033 dedicating themselves more fully to that task.
1034
1035 **P1:** Dan Schwartz example: trying to predict ultimate learning and transfer by
1036 looking at persistence through failure. Found that persistence through
1037 initial failure is the most predictive variable.
1038
1039 **P6:** Turns conversation to text and video mining.
1040
1041 **P2:** Notes the changes in the kinds of activities that are being assessed,
1042 from automated tutors to studied online learning platforms that include chat)
1043 **P8:** **Carolyn Rosé** studies talk in online multiparty conversation platforms.
1044 **P2:** A lot of the work being done focuses on scripted tasks. There is also a
1045 disconnect between the computer science folks and the learning sciences folks

1046 because places like Carnegie Mellon and other places do not have schools of
1047 Ed or Learning Sciences programs.

1048
1049 **P6:** You need a hypothesis to cut through all that data. **P2:** CS folks don't
1050 have to worry about this because they can publish something each time they
1051 identify a pattern. **P6:** Data mining - a resource that has to get used, but
1052 how? **P1:** **Joel C. Kuipers** (George Washington U linguist).

1053
1054 **P8:** This is a power issue: Data miners tend to control the discourse, tend to
1055 define what counts as evidence because they have more expertise say, than the
1056 practitioners. **P6:** Computer science community only operating at the level of
1057 algorithms - they want people to come with specific questions so they can
1058 test their algorithms TAPE ENDED AT SOME POINT HERE

1059
1060 **Day 2, Video 3**

1061
1062 **P1:** Halo videogame example: Fourth grader studies analytic of other kids
1063 coming to a Halo party so he can know how to kill them when the party begins.
1064 Evidence of expertise in analyzing complex data.

1065
1066 **P8:** Most of data mining algorithms are not community data. **P3:** The tension
1067 with the business community wanting to make this data proprietary.

1068
1069 **LUNCH BREAK**

1070
1071 **Day 2, Video 4**

1072
1073 **P8 Presentation:** Tools for Studying the Mutual Constitution of Persons and
1074 Cultural Practices in Places (Notes partially based on text from **P8's** slides)

1075
1076 Project is a collaboration with Falk, Dierking, Haun-Frank @ Oregon SU &
1077 Kirshner, York (University of Colorado) Parker (MetaScape).

1078
1079 Aim: develop approach for empowering communities to envision and create a
1080 better, more effective community-wide STEM education system. The project has
1081 two integrated components: A participatory approach to modeling and improving
1082 coordination of organizational supports for STEM learning opportunities;
1083 longitudinal study of interest development among all 5th graders in the
1084 community

1085
1086 CLAIMS: (1) Young people's sustained engagement in science-related pursuits
1087 depends on their interest levels. (2) **Tom Weisner's** (UCLA) work on **family**
1088 **routines:** How families organize their time shapes their ability/tendency to
1089 pursue interests. (3) Interests increase depending on social capital. (4)
1090 Interests depend on whether science is domain of concerted cultivation by
1091 some agents. (5) Spatial capital matters for interest development: geography;
1092 access perceived and real to resources. Spatial organization of learning

1093
1094 **P1:** SB Heath resource mapping. **P8:** Community asset mapping (Northwestern) in
1095 the context of Digital Storytelling

1096
1097 TIMELINE SLIDE - Digital storytelling and action research: 11 adolescents
1098 ethnographic informants produce digital stories, help in identifying
1099 structure of opportunities and resources for pursuing STEM learning.
1100 Digital stories inform development of surveys

1101

1102 Community meetings to identify initial model of coordinated supports.
1103 Community members are given sample agent-based models to think through that
1104 represent a psychological model but turned around: not the kid's individual
1105 interest, but how does a set of sociopsychological dynamic processes produce
1106 a distribution of kids in a community that vary in their level of interest.
1107
1108 Youth ethnographers collaborate with researchers to follow families over four
1109 years. Planned community meetings to interpret results (Community-based
1110 Intervention Research) Same project will also be mounted in east Boulder.
1111 Comparative potential.
1112
1113 One tension: no one is an insider of the community; however, Community sees
1114 this work as a service for the community which helps develop the researchers
1115 reputation as "**trusted outsiders**".
1116
1117 Social Science Constructs - the difficulty of theorizing interest.
1118 Attempt here is to think about the combination of constructs that deal both
1119 with the psychological dimensions and organizational and institutional
1120 dimensions. Psych: using **K. Ann Renneger** model of interest - dominant model.
1121 **P9**: Where is the interpersonal plane between the individual and the
1122 organizational. **P8**: Will introduce peer group dynamics later.
1123
1124 **P6**: Community level emergents? **P8**: Mutual constitution: Activities that come
1125 and go but that increase in importance as peers join them over time, compared
1126 to schools or football leagues. This will be assessed over three years.
1127
1128 **P8**: Study begins at 4th grade which is a critical time period for the
1129 development of interest in science (Kids all say they're interested in
1130 science at the beginning but by the end of the year only a quarter claim
1131 interest). **P6**: Competing interests (as instantiated in the agent-based
1132 model)? **P1**: Will have to be assessed qualitatively, through case studies.
1133 **P6**: Competing interests that could be linked or transformable to STEM, but
1134 others that actually draw people away from STEM. **P2**: In STEM you need some
1135 kind of hardware to do things - not just digital tools which can bundle
1136 everything - there is something important about the physicality.
1137
1138 **P8**: The focus in this intervention is on coordination to help people in the
1139 community recognize local community resources.
1140
1141 **P2**: STEM hardware tends to be fixed in a location. His group is trying make
1142 mobile tools for kids to take home.
1143
1144 **P1**: Issue of aquarium positioning kids of different ethnic groups - the
1145 Anglos do the prestige science work, the Latino kids are given "on the floor
1146 job work". How would you see this adverse positioning in this project? **P1**:
1147 Alerting the youth ethnographers to this.
1148
1149 **P9**: Asks about similarity of this project to Engestrom's Change Labs - there
1150 are no institutional contradictions here. **P8**: Community meetings component
1151 will address this because they will externalize the production of inequality,
1152 of declining interest. **P9**: Becker "see how an institution produces its
1153 failures".
1154
1155 Agent based models slide **P8**: What is of interest here is that the models will
1156 be put in the hands of the community as tools for thinking
1157

1158 Initial model: Social resources slide. Guides have an impact on learners by
1159 encouraging their interest in learning itself. Not all guides are created
1160 equal: some guides are much more effective at increasing learning interest
1161 over time. No matter how effective a guide is, there are limits to how much
1162 effect they can have on a given learner. Learner's learning interest grows by
1163 itself over time. This effect is self-reinforcing, the more a learner becomes
1164 engaged in learning, the more their learning interest will grow. There is a
1165 minimum level below which some learners will not be able to grow their own
1166 learning interest at all, and this is where a guide's influence could be most
1167 crucial.

1168
1169 Modeling interest development Slide. Provides visualization of interest
1170 development with respect to geographic location, movement constraints,
1171 resource availability. P3: Resources in the home? P8: Not included yet.

1172
1173 P7 LEAVES

1174
1175 P9: Questions assumptions about intrinsic motivation in this model. Why not
1176 include other people as a scaffold for interest?

1177
1178 P3: With spatial models we also need to think about virtual spaces.
1179 What does STEM mean to the kid? One thing is for the learner to be interested
1180 in scientific things, it's another to develop an allegiance to evidence. P8:
1181 Youth ethnographers are being recruited to identify topics of interest.

1182 P3: What if a kid says he really likes Second Life. Does that count as STEM
1183 interest? P8: Haven't worked this out yet. P8: What are good markers that a
1184 kid can respond to on a large-scale survey? You can't do a mapping of a tool
1185 to a practice because a tool can be used to support multiple practices
1186 P6: What constitutes a stem interest may change. For example, the question of
1187 connection between second life and geometry. P9: You don't want to reproduce
1188 the normative understandings of STEM in order to get people to like them.

1189 P1: You can look at non-science social domains and show the participants the
1190 points of connection (via structured interviews) P8: Use surveys to hone in
1191 on pursuits that participants care about most and identify markers of these
1192 pursuits.

1193
1194 P2: Discusses agent-based model research in a classroom where kids are engage
1195 in math problem solving to understand why high achieving students were given
1196 more difficult sub-tasks as compared to the low achieving kids. This points
1197 to the issue of immediate gratification and length of task (higher
1198 opportunities for disengagement)

1199
1200 P6: How to build a space for assessing how far people have gone in pursuing
1201 science interest: building a long term trajectory vs. the low-hanging fruit.
1202 P8: Rarely is the trajectory one in which there is deepening within the same
1203 activity. Azevedo, Flavio "lines of practice": you have an interest and you
1204 find different activities in which can plug it in, making it seem continuous.
1205 So you need something at a higher level that binds it all together.

1206
1207 P1: How to model the resonance of predominant resources in relation to
1208 practices in the cultural history of the group being studied in P8's work.

1209
1210 P8: A challenge here is that we want fragments of models for people to work
1211 with and that allow them to see dynamics. ABM should be used as a heuristic
1212 to help you think.

1213
1214

1215 **P6** moves over to discussion of recommendations
1216
1217 **P8**: Make it about people, knowing WHO to approach for advice for doing the
1218 work. **P9**: MacArthur should fund some set of people with sophisticated
1219 theories of learning to think about the assessment issue **P3**: This has been
1220 done: Valerie Shute assessing Quest to Learn. This example reveals "Value
1221 Differences": Shute evaluating Katie Salen - each has different values about
1222 what to assess. There needs to be more multiplicity of theories.
1223
1224 **P1**: Another model is a committee of visitors that spends some time in the
1225 program and then makes recommendations. **P3**: When you have to design an ECD
1226 model vs. Katie Salen actually having to implement a design.
1227 Example of Gamestar mechanics. This was developed as part of research but
1228 them commercialized. The commercialization brings in a new set of values.
1229 Then there is the tension that the commercialization model seems to be the
1230 only sustainable one. MacArthur is uncomfortable with this moment of opening
1231 up the research to let other value systems in. E.G. not just commercial
1232 interest, but voices from the learning sciences have been absent from the
1233 MacArthur table.
1234
1235 **P3**: The idea of interest-driven learning has lived out its time (like
1236 community of practice). We are at a stage where we're taking a qualitative
1237 model -like interest-driven or geeking out - and instantiating this and
1238 these are theories that are "way too vague" and this risks collapse. Some of
1239 the underlying claims here are suspect. When you come in and do the
1240 assessment, this is the moment when you see how they are suspect because now
1241 you have to provide evidence.
1242
1243 **P9**: You want evaluators that have some sense of the endogenous forms of
1244 assessment and that can translate this to the outside world. **P6**: e.g. Four
1245 funded projects each with house assessors to motivate assessment from the
1246 very beginning. At intervals, the other three assessors visit with the one to
1247 confer about assessments. At the end the four assessors write the report
1248 about what they learned to do about doing assessment across the four
1249 different projects. **P3**: We need good examples, case studies of where the
1250 thing worked. The literature review for the report reveals how disparate the
1251 efforts have been. **P6**: This four case model might be a good way to turn some
1252 of the differences that can disrupt good evaluation within a project into an
1253 asset of getting an understanding evaluation approaches across projects. The
1254 four examples learn from each other.
1255
1256 **P1**: NRC report may be worth revisiting. DBIR as a method should be proposed
1257 as a recommendation.
1258
1259 **P8**: Computer science work examining design rationales. Looking early on at
1260 what counts as evidence. Connected to Toulman (like ECD)
1261
1262 **P3**: Example of game designers keeping diaries. This could be applied in DBIR.
1263
1264 **P2**: There should be initiatives to educate staff in afterschool programs how
1265 to interpret and use assessment tools.
1266
1267 **P1**: What terms to use: assessment, evaluation, research. **P8**: Pamela Moss -
1268 different kinds of assessment: discursive, inherent, documentary (coming out
1269 of ethnographic research) **P2**: Need for programs to train graduate students in
1270 these methods.

1271 **P3**: This is an interdisciplinary endeavor. Need for a fellows program for
1272 training grads to do this interdisciplinary work.
1273
1274 **P8**: Need to connect after-school corporations in these conversations (e.g.
1275 TASC) because they are central to capacity building. We need to network with
1276 youth development organizations.
1277
1278 **P3**: The trouble with B&G clubs is that the kids are dropped in so you can't
1279 do longitudinal work. You always have different kids
1280
1281 **P1**: One point of consensus from the conversation yesterday was the focus on
1282 **engagement**. **P9**: Engagement means many things. Perhaps developing a field
1283 guide for explaining the varieties of engagement, what are the different
1284 markers of engagement
1285
1286 **P2**: More work examining the quality of the work done by kids in intensive
1287 projects like **P2**'s and connecting it to the standards of the profession. The
1288 concern here is with the attitude of "I just want kids to make stuff"
1289 **P8**: In contrast to this: Youth Radio, Lisa Soep. Does interesting work on
1290 professional level production.
1291
1292 **P3**: Choice is crucial in after school, but then you move choice into school
1293 and it means something totally different. We need a theory of the essential
1294 features that make things work.
1295
1296 **P8**: It is important to consider approximation of professional practices. **P3**:
1297 And the development of internal high standards that the community itself
1298 polices, negotiates. **P6**: This takes time. We not only have to consider
1299 reasons of equity but also of affect (e.g. you can't just critique the
1300 learner from day one).
1301
1302 **P3**: A principle in game design is that you have to have a success in the
1303 first two minutes. **P6**: One problem lies in short programs where your ability
1304 to critique is really constrained because you simply don't have the time to
1305 slowly bring the kid up to speed without discouraging him.
1306
1307 **P8**: Study how do standards emerge for quality. "The emergence of critical
1308 standards" Richard Lehrer & Julie Erickson. **P3**: The difficulty of community
1309 building consensus of criteria, and the problem of getting the kids to agree
1310 and meet these standards without discouraging them. **P6**: Writing community is
1311 a good example of consensus **P8**: They have a long history of social practices
1312 where you expect to be critiqued.
1313
1314 **P3**: There are intermediate levels, so different people have different levels
1315 of status **P9**: Dissertation work on architectural critique. The critiques are
1316 so eviscerating that **P9** had to come in and act as a translator-buffer between
1317 architects and kids who participated in the program.
1318
1319 **P6**: Model of enlarging communities: Doing well compared another kid, vs. to a
1320 group, vs. to science community **P3**: e.g. Cohort matching in video games as a
1321 practice in videogame communities. **P8**: Roger Hart Ladder of Participation
1322
1323 **P6**: BADGES as a way to make this visible. **P8**: Is concerned about badges
1324 because they depend on a system of recognition that is institutionalized. We
1325 need systems of recognition that go across institutions. **P3**: Badges have the
1326 advantage of allowing one to map out the geography of where you're going.

1327 P6: Levels of aspiration. P8: Badges need people who do the hand off - a
1328 broker who helps you navigate. P3: This is crucial for Macarthur - they need
1329 to know this: brokers, ladders of participation. P6: What if every badge had
1330 two levels - achieving it and teaching others how to achieve it. P1: Nichole
1331 Pinkard: community level badges. P2: Concern: Badges might become so
1332 competitive that kids for example might not join robotics teams because their
1333 too competitive.
1334
1335 P3 & P5 Leave.
1336
1337 P1: The risk of the out of school environments having to connect badges to
1338 things that they are accountable to in school. But is this bad? Need to think
1339 critically about this P9: What if badges disrupt the old assessment economy?
1340 P6: Building in altruism into the badge? P8: Badges represent qualities of
1341 achievement P9: Bourdieu argument: badges will jut translate into the
1342 universal capital of grades P8: How can one work against the entire network
1343 of schools organizing the testing juggernaut. P1: Washington state has
1344 alternatives to testing that could line up with badges as a good counterpoint
1345 to standardized testing. P8: Badges as a means of broadening participation.
1346 So one consensus position would be a careful study of how badges develop
1347 within the different communities. P9: Look at the relationship between the
1348 badge economy and the real economy P2: We need to study the relationship
1349 between badges of traditional content, and badges for non-traditional
1350 activities. P9: We have to keep in mind the question of endorsement.
1351
1352 P6: What about portfolio assessment? Then you don't have one common standard
1353 of assessment.
1354
1355 P8: Key Empirical question: To what extent can badges remain local and if so,
1356 what is the boundary. P6: National scale endorsement of badges would turn
1357 into standardized curricula and assessables.