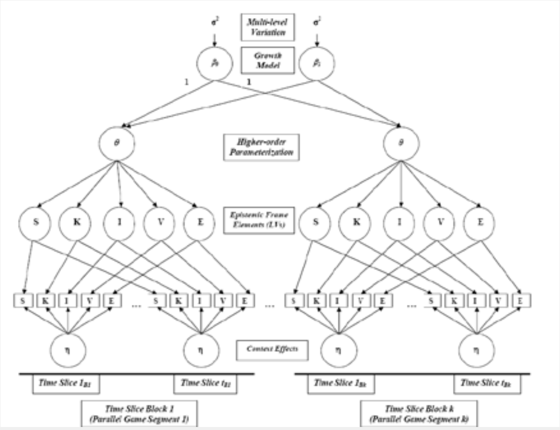
Developing, Modeling, and Representing Domain-specific Expertise via Epistemic Games: A Measurement Person's View

André A. Rupp
EDMS Department
University of Maryland
1230-A Benjamin Building
College Park, MD 20742

E-mail: ruppandr@umd.edu
Phone: (301) 405 – 3623
Fax: (301) 314 – 9245



Physics Education Research Group
University of Maryland, February 22, 2010

Genres of Games

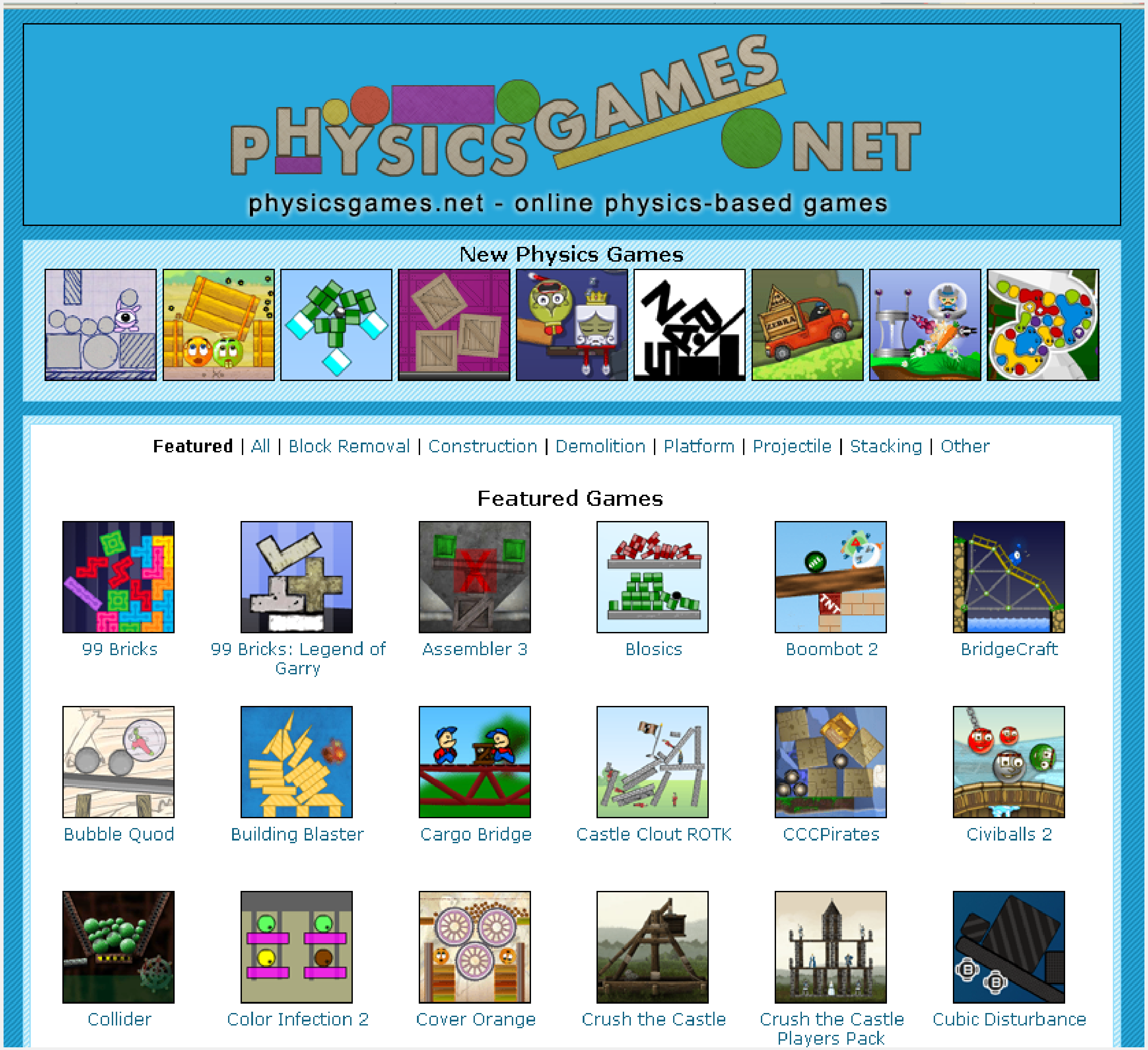
1. Nature of learning connected to game
    a. Develop inquiry / argumentation skills
    b. Develop expertise characterizing target community
    c. Content and process learning (non-simulation based)
    d. Content and process learning (simulation based)
    e. Familiarity with alternative representations, tools, and processes

2. Duration and nature of game participation
    a. Short interaction / casual games
    b. Longer duration finite games
    c. On-going participation games

3. Intended purpose of game
    a. Fully recreational games for entertainment purposes
    b. Serious games for informal learning
    c. Serious games for formal learning
    d. Assessment games

adapted from Clark, Nelson, Sengupta, & D'Angelo (2009)

Entertaining Simulations / Mini-games



Educational Simulations / Intelligent Tutoring Environments

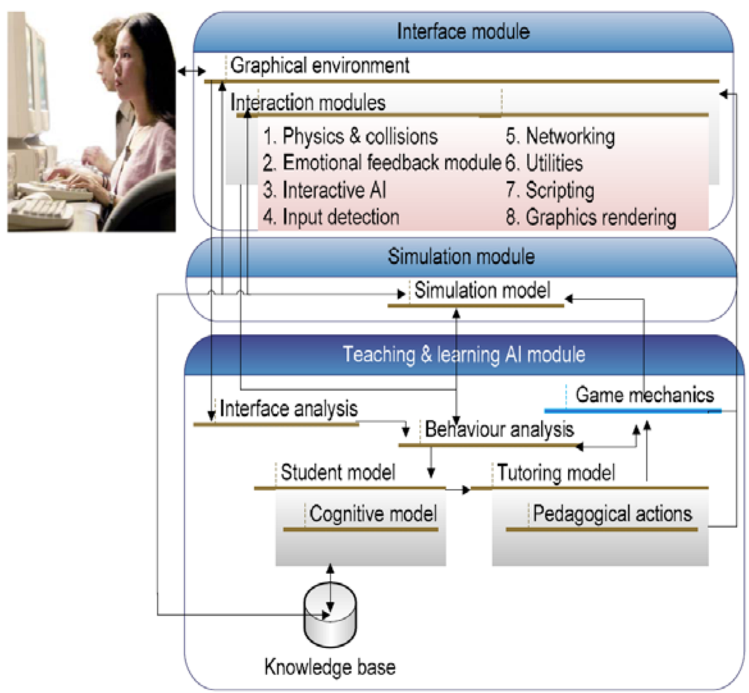
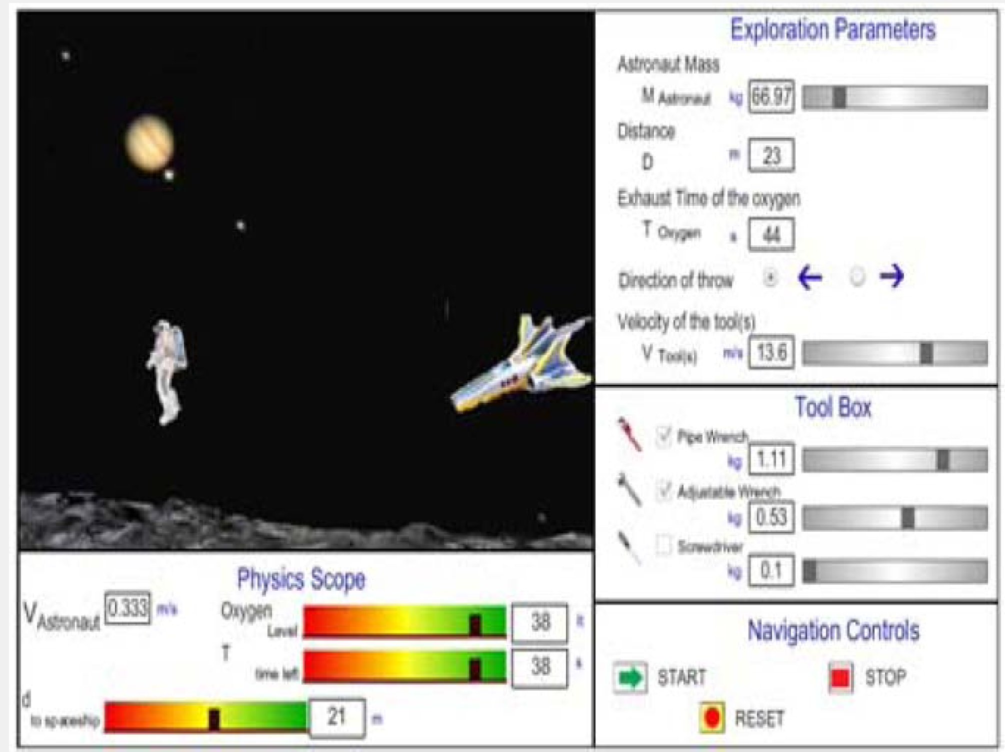
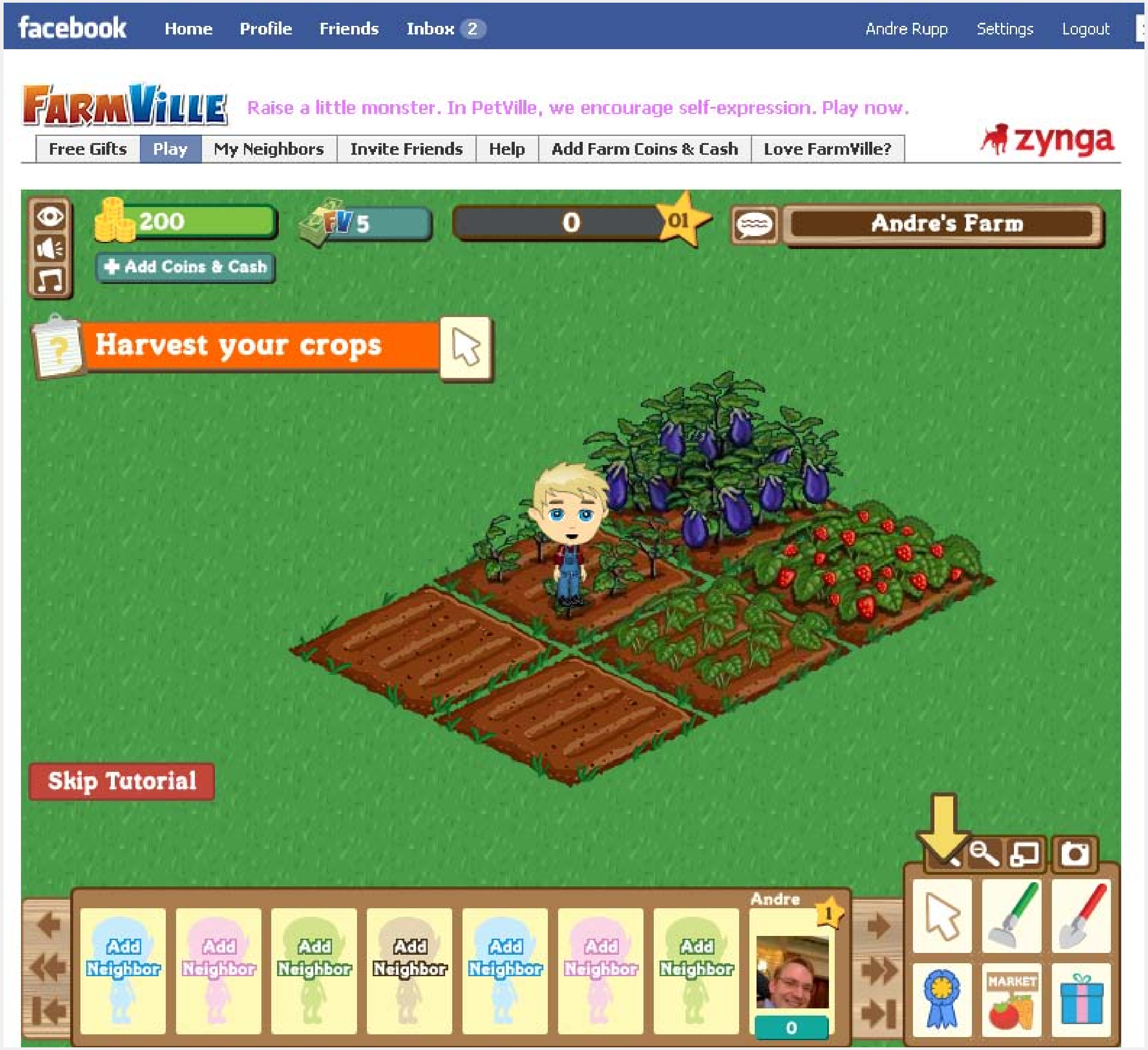


FIGURE 1
OLYMPIA ARCHITECTURE



Muños et al. (2009)
http://fie-conference.org/fie2009/papers/1457.pdf

Massively Multiplayer Online Games



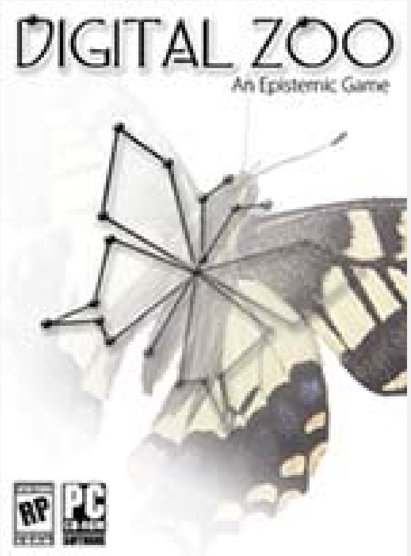
http://apps.facebook.com/onthefarm/gifts.php?ref=interstitial

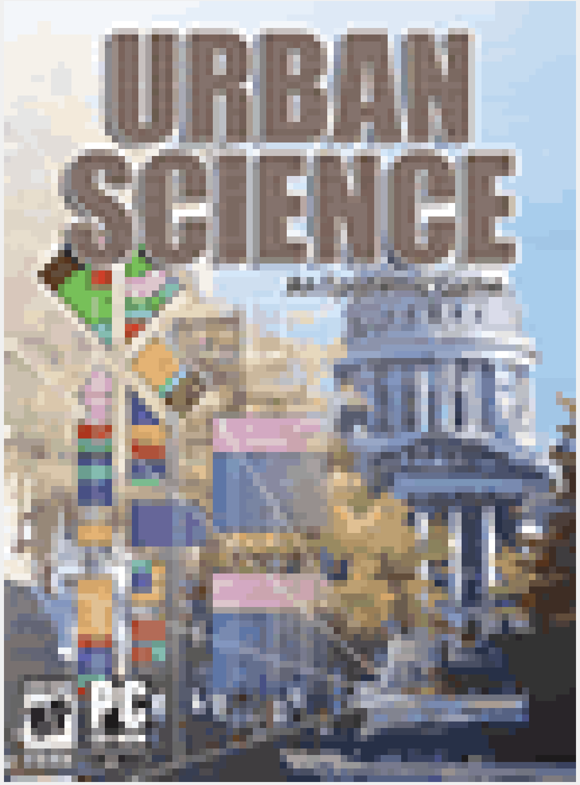
Virtual Worlds: Quest Atlantis

	A	B	C
Turbidity	6	**27**	22
Dissolved Oxygen	5.5	**4.5**	4.0
Temperature	17.5	**22.5**	22.0
Nitrates	3.15	0.96	2.08
Phosphates	3.6	1.7	3.1
pH	6.6	7.0	7.3

http://atlantis.crlt.indiana.edu/

Collaborative Platforms: Epistemic Games





science.net

www.epistemicgames.org

Objectives of Epistemic Games

To help players think and act like real-world professionals in a discipline by…

… presenting them with realistic and meaningful problems along with suitable constraints from that discipline.

…providing them with scaffolded support from trained mentors who encourage them to develop such disciplinary reasoning and action.

…encouraging them to make numerous links between different

types of skills (S),
types of knowledge (K),
aspects of their identities (I),
aspects of their value systems (V), and
aspects of their epistemological beliefs (E).

These SKIVE elements are intimately interconnected in the epistemic frame of players.

UDA Mail: Andre

Reply Close

From	Maggie
Sent	Friday, December 4th, 2009 at 9:41 am
To	Andre
Subject	Welcome!

Hi Andre, it is nice to meet you. My name is Maggie. I've been working here for about five years as the community facilitator.

I am the contact person between the firm and community. I will help you learn how to fulfill your responsibilities as a planner, too. The other person who will be helping you is your planning consultant. You should get in touch with your planning consultant by pressing the "chat" button on the right. Always leave your chat window open, so that you can communicate regularly.

The first thing you need to do is answer some questions for our intake interview. Don't worry about whether you know the answers; just do the best you can. Your answers will help us improve the experience of new planners like yourself. You can get started **here**.

Welcome!
-Maggie

Inbox
- Sent Msgs

Planning Notebook

Chat

Project: State Street

Phase 1: *People for Greenspace*
- Virtual Site Visit
- iPlan: Practice map
- Zoning Matrix
- iPlan: Preference Survey

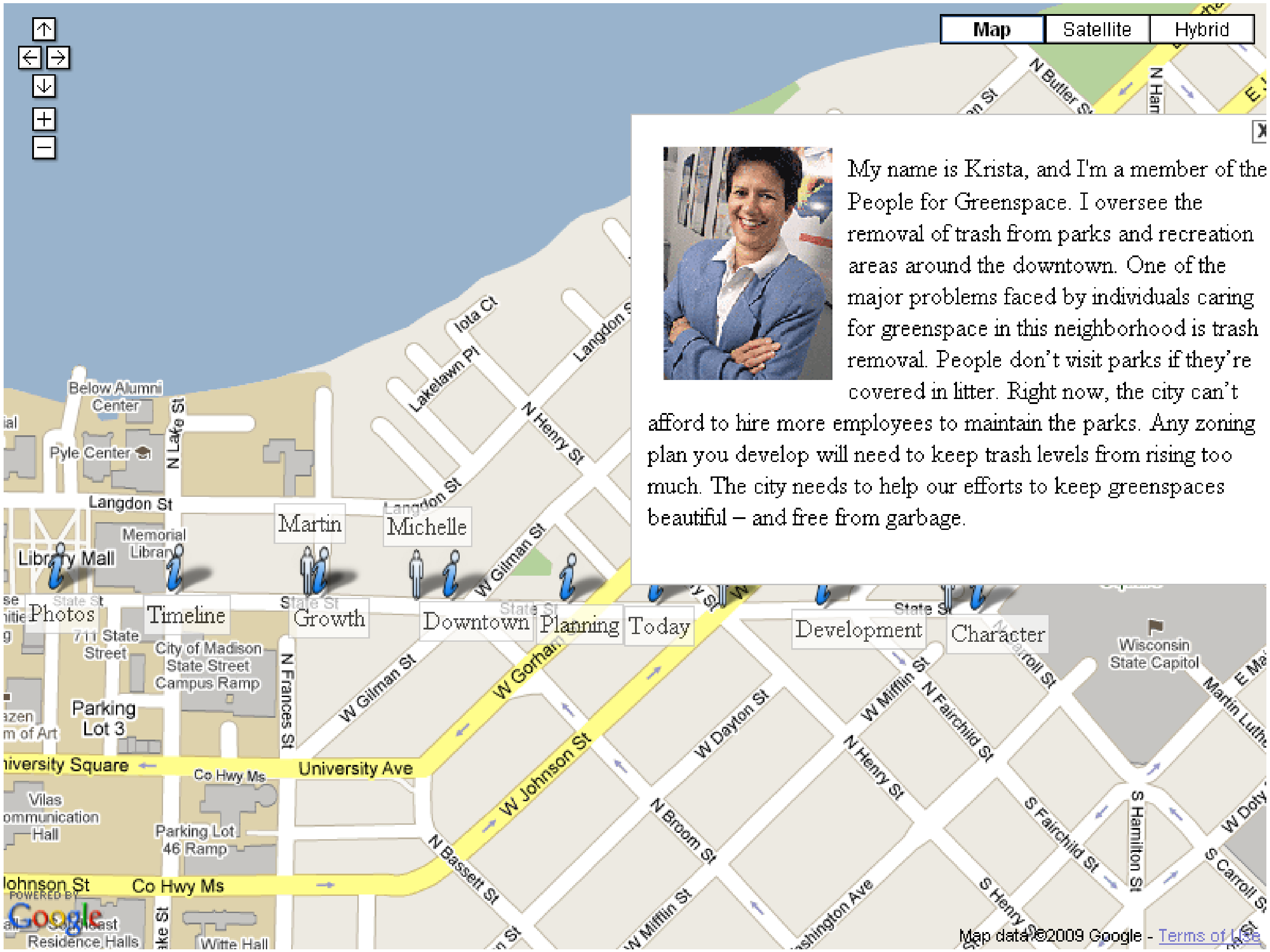
Urban Design Associates
- Staff
- Professional Resources

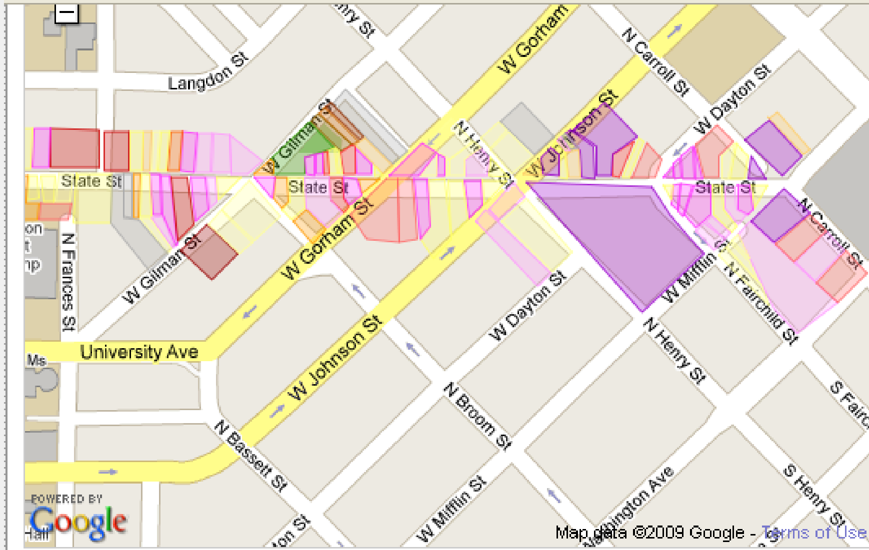
Logout Andre

http://epistemicgames.org/demo_us8/?page_id=24

Virtual Site Visit

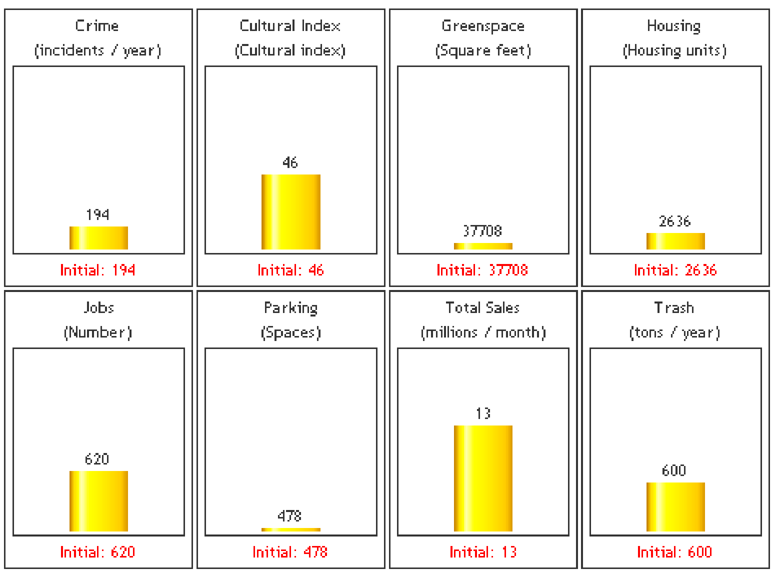
To visit stakeholders or review background information, select the appropriate icon in the map.





Preference Survey Matrix Targets

Cultural Index: 50 Parking: 200 Greenspace: 40000 Trash: 650



AR	Arts and humanities
C1-L	Local retail store
C1-L-R3	Local retail store with low density housing
C1-L-R4	Local retail store with high density housing
C1-N	National chain retail store
C1-N-R3	National chain retail store with low density housing
C1-N-R4	National chain retail store with high density housing
C2-L	Local restaurant
C2-L-R3	Local restaurant with low density housing
C2-L-R4	Local restaurant with high density housing
C2-N	National chain restaurant
C2-N-R3	National chain restaurant with low density housing
C2-N-R4	National chain restaurant with high density housing
OS	Green and open space
P-G	Parking garage
P-S	Surface parking
R1	Single family home
R2	Housing with 1-4 units
R3	Housing with 4-8 units
R4	Housing with more than 8 units

Stakeholder Assessment

Cultural Index | Parking | Greenspace | Trash

Initial value	**37708 ft^2**			
		Stakeholder Feedback	**Acceptable Max/Min**	**Unacceptable Max/Min**
Andre	40612.21 ft^2			
Brent	173780.26 ft^2			
Kathy	74877.79 ft^2	Good		

Update

Construct Labels for Epistemic Frame

Skills (various): being able to communicate clearly, both orally and in writing; being able to collect, organize, and analyze information; being able to think critically and justify different positions; being able to view issues from the perspective of others.

Knowledge (terms of art, systems thinking): knowing institutions and processes that drive civic, political and economic decisions; knowing how a community operates, the problems it faces, and the richness of diversity.

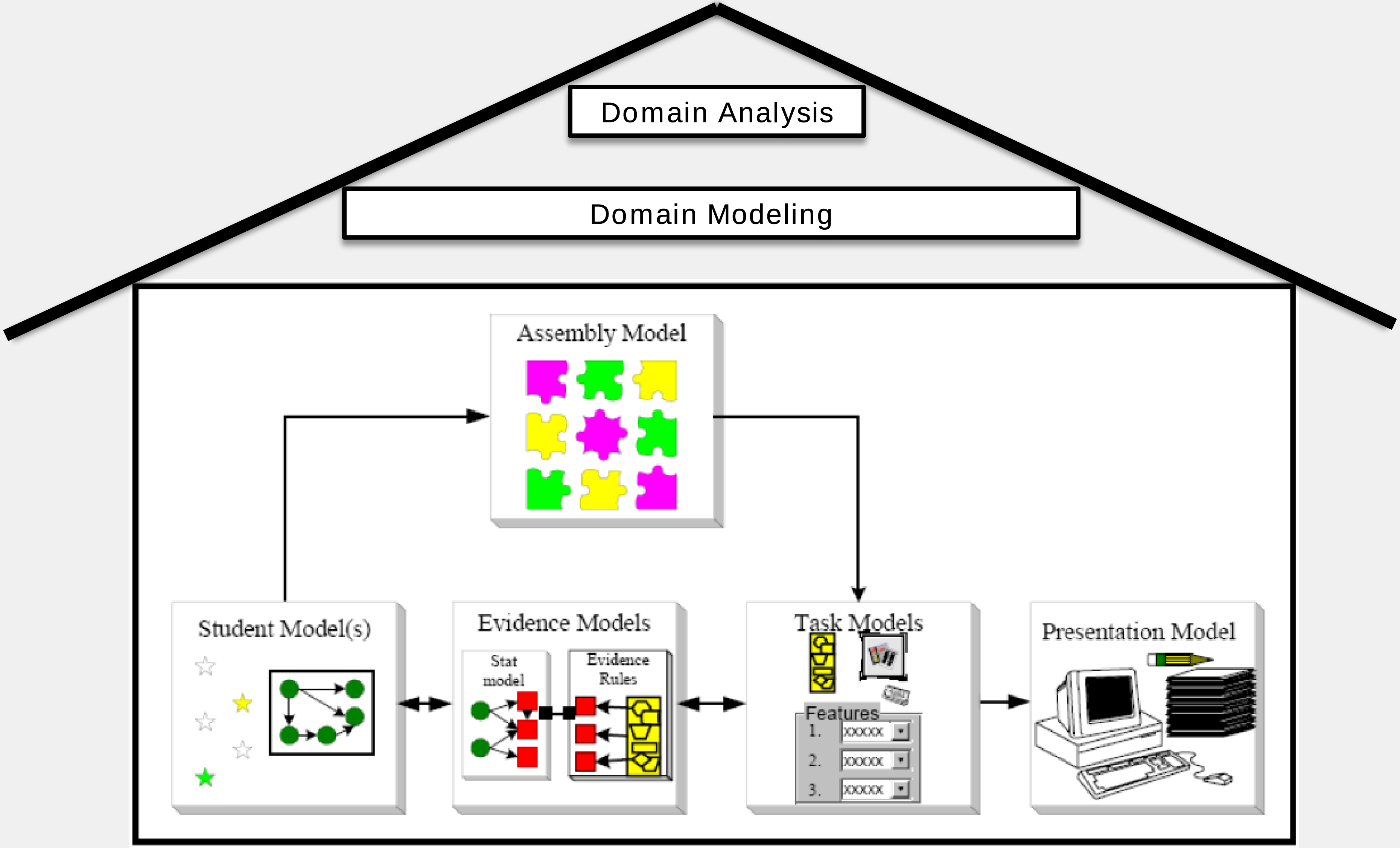
Identity (as planner, as professional): having a way of seeing oneself that is commensurate with how members of the urban planning community see themselves.

Values (working for stakeholders, for the public good, as a team, like a professional): being willing to listen to, and take seriously, the ideas of others.

Epistemology (general, planning-specific): being able to understand what counts as relevant evidence that justifies actions as legitimate within the urban planning community.

Linking Epistemic Games Design and Resulting Data Structures

Evidence-centered Design



ECD & Validation Efforts

1. Content validity
   ⇒ does the content of the assessment represent the target domain?
2. Substantive validity
   ⇒ do the respondents engage in the appropriate cognitive processes?
3. Structural validity
   ⇒ does the scoring process reflect the interaction of abilities in the domain?
4. Predictive validity
   ⇒ can the assessment scores be used to predict an outcome of interest?
5. External validity
   ⇒ do respondents perform similar on assessments tapping similar constructs and differently on assessments tapping different constructs?
6. Generalizability
   ⇒ can the assessment results be generalized across different conditions such as time points, administration contexts, and respondent samples?
7. Consequential validity
   ⇒ do the assessment interpretations lead to fair and defensible consequences for respondents?

Epistemic Frame Representations

Time Slice	S	K	I	V	E
1	1	1	0	0	0
2	0	1	0	0	0
3	0	0	0	0	0
4	1	1	1	0	1
:	:	:	:	:	:
T	1	1	0	1	1

Illustration of Evidence Accumulation in ENA

Time Slice	S	K	I	V	E
1	1	1	0	0	0
2	0	1	0	0	0
3	0	0	0	0	0
4	1	1	1	0	1
:	:	:	:	:	:
T	1	1	0	1	1

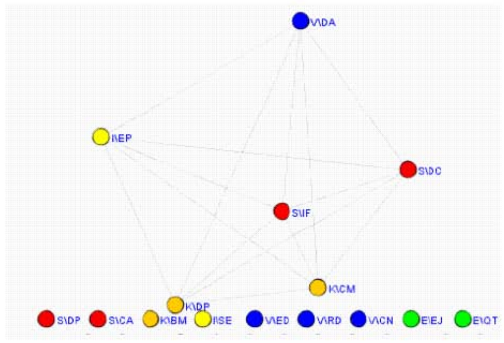
	S	K	I	V	E
S	0	1	0	0	0
K	1	0	0	0	0
I	0	0	0	0	0
V	0	0	0	0	0
E	0	0	0	0	0

	S	K	I	V	E
S	0	1	1	0	1
K	1	0	1	0	1
I	1	1	0	0	1
V	0	0	0	0	0
E	1	1	1	0	0

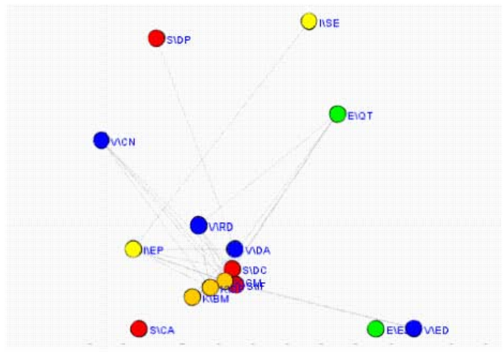
At end of game play

(or at particular time slice t)

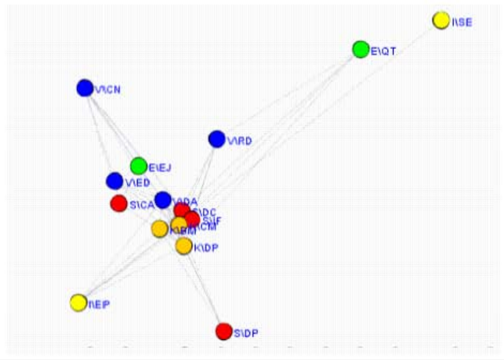
	S	K	I	V	E
S	0	1	14	7	9
K	23	0	11	3	8
I	14	11	0	16	15
V	7	3	16	0	13
E	9	8	15	13	0



Early Stage
Of Game Play



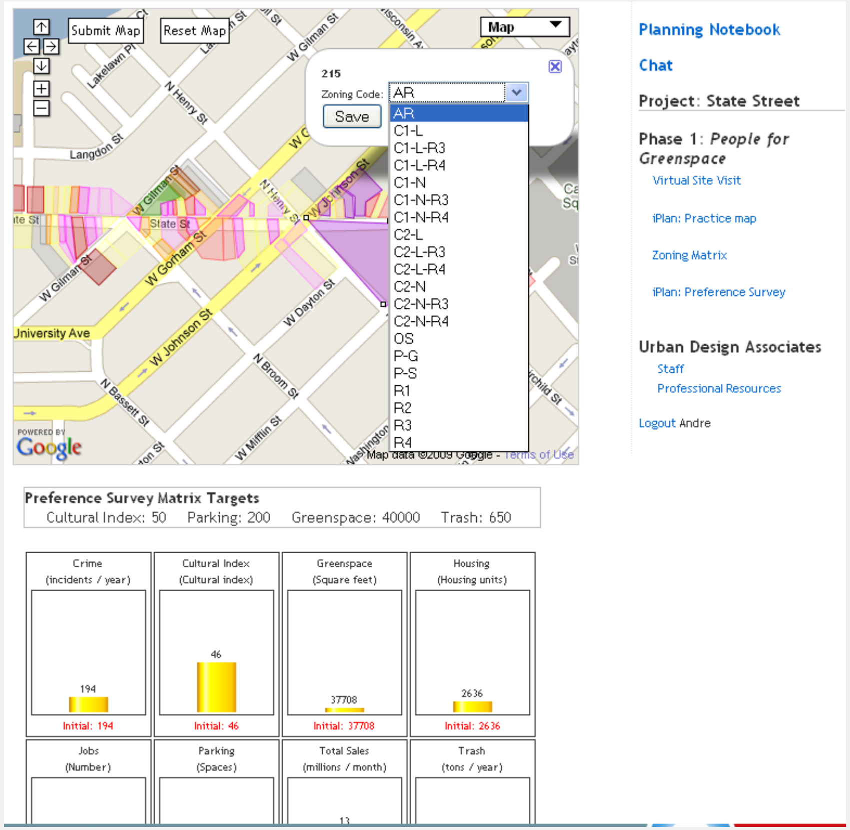
Middle Stage
Of Game Play



Late Stage
Of Game Play

Properties of Resulting Data Structures

1. **Multivariate nature of data**
2. **Discrete nature of codes**
3. **Longitudinal nature of data**
4. **Context-dependency of data**
5. **Association focus of interpretations**
6. **Intra-individual focus of interpretations**



Measurement Objectives

Statistical models for data from epistemic games are used to…

… operationalize the frame elements

… represent their association structure and individual states

… model the development of these structures longitudinally

… represent different models of learning

… compare different epistemic frame representations

… quantify uncertainty associated with inferences

One particularly promising analytic method is epistemic network analysis but other latent-variable models (BINs, DCMs, HMMs) may be promising too. What are their respective advantages and disadvantages?

ENA Simulation Study

Investigations into ENA

Idea

Generate data that have the same structure as observed data

Use ideas from latent-variable models to do so, specifically the separation of task and respondent parameters

Postulate patterns for task and respondent parameters, generate data, and see whether ENA statistics are sensitive to variations in these parameters

Simulation Study

About 50 conditions for learning progressions

About 10 conditions for task parameters

Various outcome statistics

Mathematical Idea

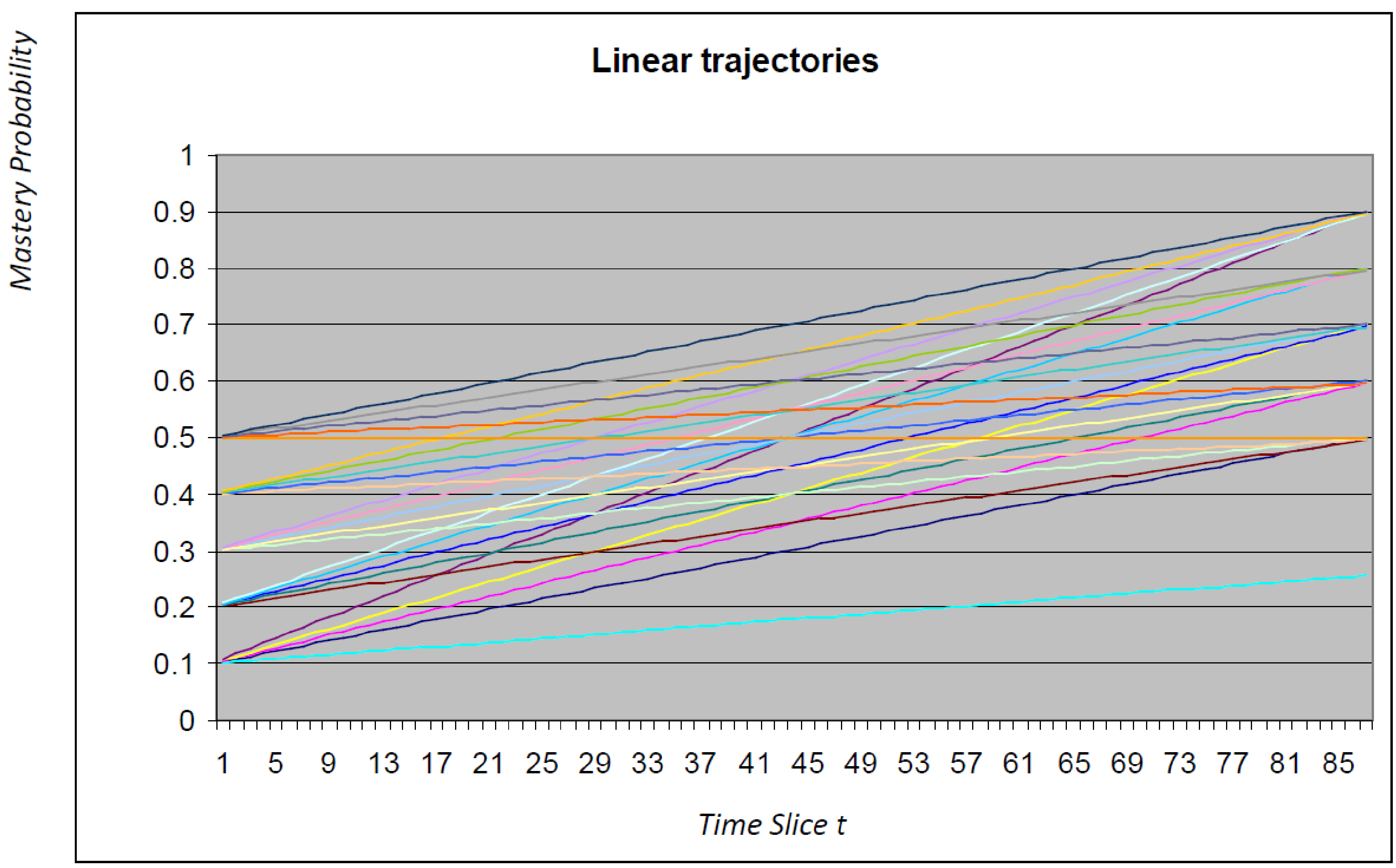
Conditional Probabilities

Mastery	Non-mastery
$P(X = 1 \mid \alpha = 1, E = 1)$	$P(X = 1 \mid \alpha = 0, E = 1)$
$P(X = 0 \mid \alpha = 1, E = 1)$	$P(X = 0 \mid \alpha = 0, E = 1)$
$P(X = 1 \mid \alpha = 1, E = 0)$	$P(X = 1 \mid \alpha = 0, E = 0)$
$P(X = 0 \mid \alpha = 1, E = 0)$	$P(X = 0 \mid \alpha = 0, E = 0)$

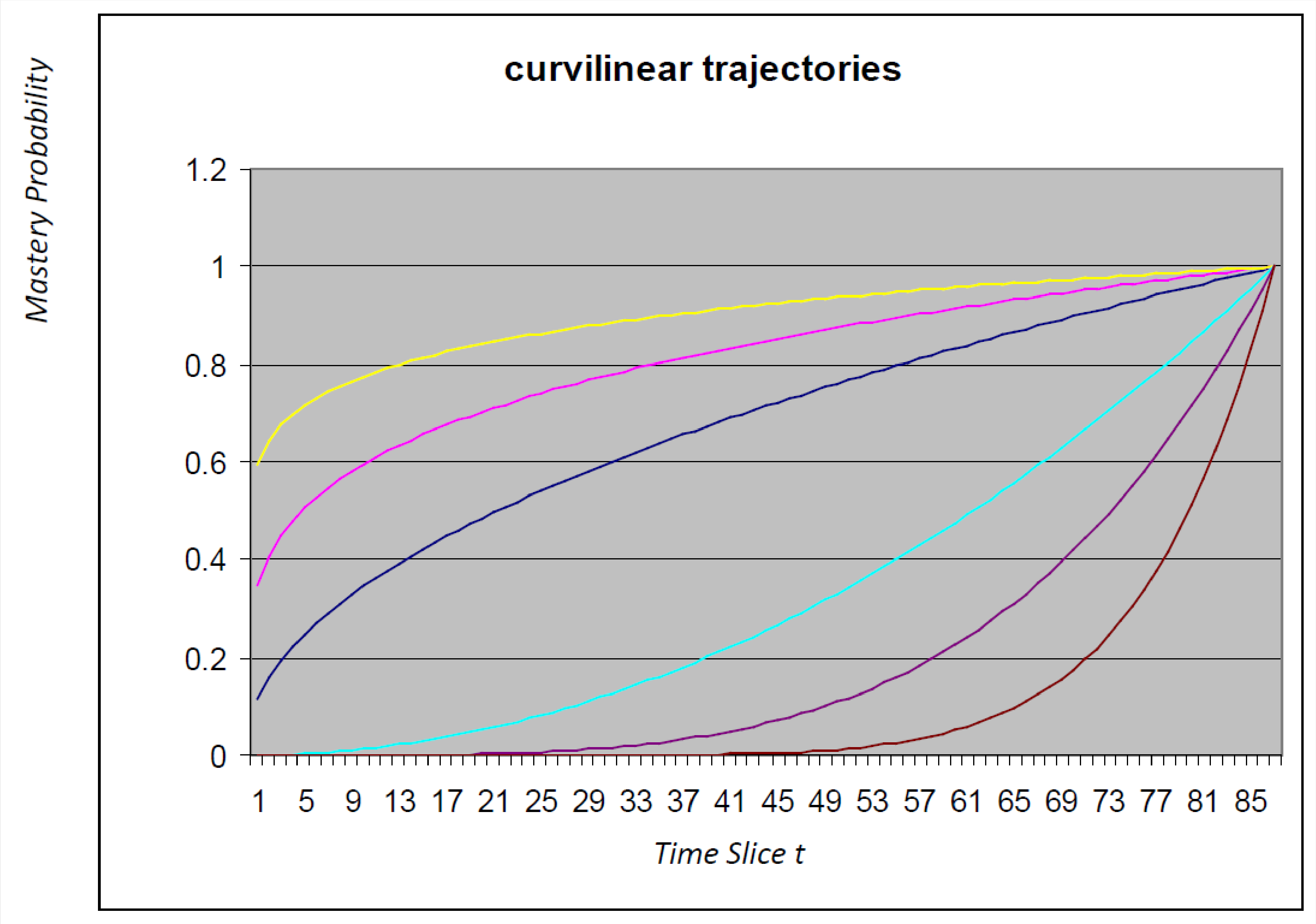
Marginal Probabilities

$$P(X = 1 \mid E = 1) = P(X = 1 \mid \alpha = 1, E = 1)P(\alpha = 1) + P(X = 1 \mid \alpha = 0, E = 1)P(\alpha = 0)$$
$$P(X = 0 \mid E = 0) = P(X = 0 \mid \alpha = 1, E = 0)P(\alpha = 1) + P(X = 0 \mid \alpha = 0, E = 0)P(\alpha = 0)$$

Learning Trajectories (Non-linear Trends)



Learning Trajectories (Non-linear Trends)



Task Parameter Settings

$$P(X=1|\alpha=1,E=1)=1-s^{(1)}$$
$$P(X=1|\alpha=0,E=1)=g^{(1)}$$
$$P(X=0|\alpha=1,E=0)=1-s^{(0)}$$
$$P(X=0|\alpha=0,E=0)=g^{(0)}$$
$$P(\alpha=1)=P(\alpha)$$

Condition	S		K		I		V		E		Description
	s(1)	g(1)	s(1)	g(1)	s(1)	g(1)	s(1)	g(1)	s(1)	g(1)	
1	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	All tasks are well-designed
2	High	High	High	High	High	High	High	High	High	High	Tasks are poorly designed
3	Low	High	Low	High	Low	High	Low	High	Low	High	All tasks are slightly easy, it's easy to guess the answers, and it's also it's difficult to fail to demonstrate a skill
4	High	Low	High	Low	High	Low	High	Low	High	Low	All tasks are slightly difficult; it's difficult to guess the answers, and it's not hard to fail to demonstrate a skill when it is called for
5	Low	High	Low	High	High	Low	High	Low	High	Low	Easy to demonstrate the low level skills, Difficult to demonstrate the higher level skills
6	Low	Low	Low	Low	High	Low	High	Low	High	Low	Well-designed tasks to elicit the demonstration of low-level (more concrete) skills, difficult to demonstrate higher level skills
7	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Easy to demonstrate lower-level skills and tasks for higher-level skills are well-designed

Condition	S		K		I		V		E		Description
	s(0)	g(0)	s(0)	g(0)	s(0)	g(0)	s(0)	g(0)	s(0)	g(0)	
A	Low	1	Low	1	Low	1	Low	1	Low	1	Tasks are highly constrained, providing learners with few opportunities to demonstrate abilities not critical to completion of the task; probability of a novice failing to demonstrate a skill when it's not critical to the completion of the task is set to 1
B	Medium	1	Medium	1	Medium	1	Medium	1	Medium	1	Tasks provide learners with moderate opportunities to demonstrate abilities not critical to task completion; probability of a novice failing to demonstrate a skill when it's not critical to the completion of the task is set to 1
C	High	1	High	1	High	1	High	1	High	1	Tasks are open-ended, providing learners with ample opportunities to demonstrate abilities not critical to task completion; probability of a novice failing to demonstrate a skill when it's not critical to the completion of the task is set to 1

Task Parameter Settings (Part II)

Condition	S		K		I		V		E	
	s(1)	*g(1)*	*s(1)*	*g(1)*	*s(1)*	*g(1)*	*s(1)*	*g(1)*	*s(1)*	*g(1)*
1	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2	High	High	High	High	High	High	High	High	High	High
3	Low	High	Low	High	Low	High	Low	High	Low	High
4	High	Low	High	Low	High	Low	High	Low	High	Low
5	Low	High	Low	High	High	Low	High	Low	High	Low
6	Low	Low	Low	Low	High	Low	High	Low	High	Low
7	Low	High	Low	High	Low	Low	Low	Low	Low	Low

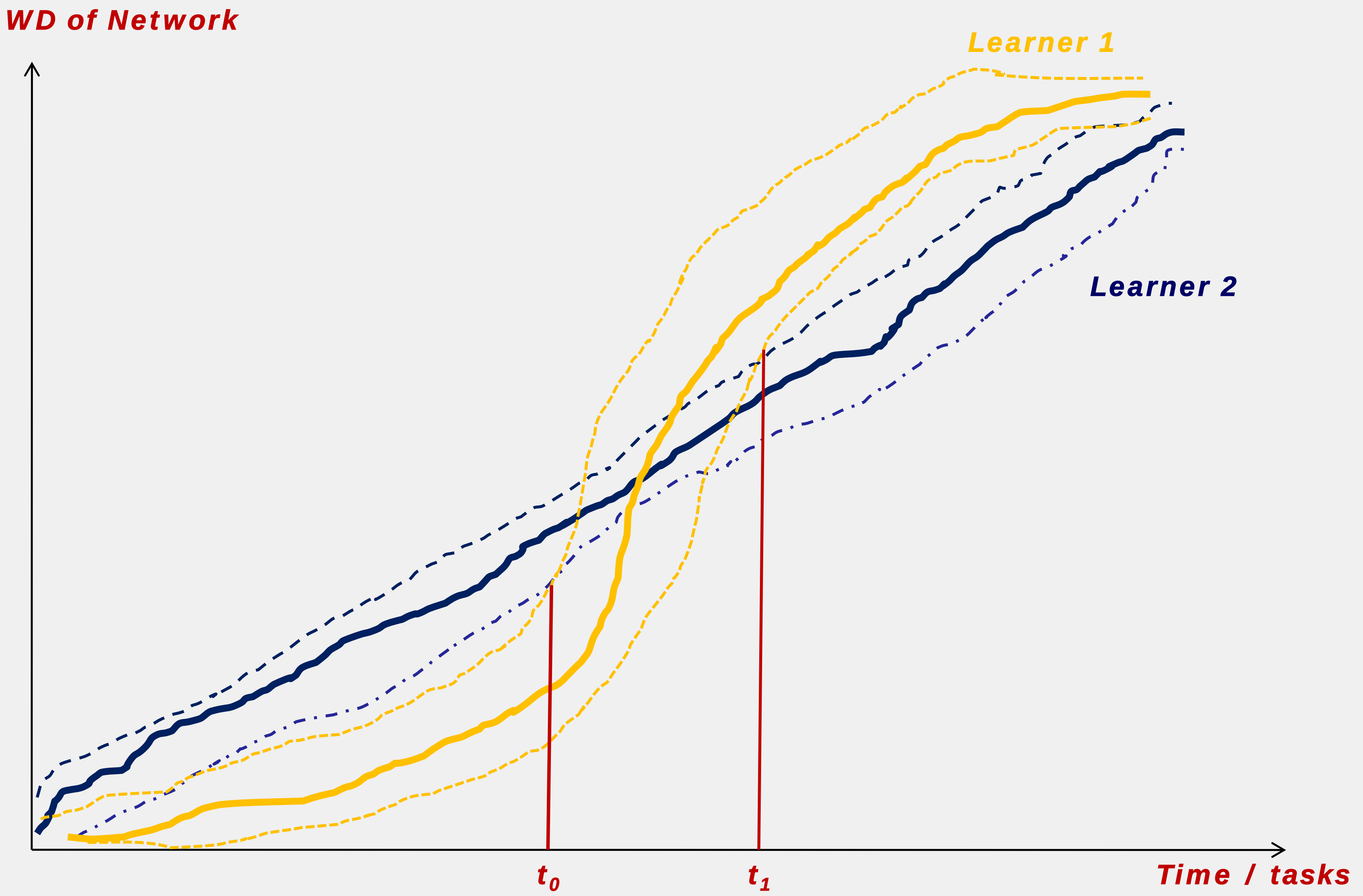
Condition	S		K		I		V		E	
	s(0)	*g(0)*	*s(0)*	*g(0)*	*s(0)*	*g(0)*	*s(0)*	*g(0)*	*s(0)*	*g(0)*
A	Low	1	Low	1	Low	1	Low	1	Low	1
B	Medium	1	Medium	1	Medium	1	Medium	1	Medium	1
C	High	1	High	1	High	1	High	1	High	1

Task Parameter Meaning

-	*Description*
	All tasks are well-designed
	Tasks are poorly designed
	All tasks are slightly easy, it's easy to guess the answers, and it's also it's difficult to fail to demonstrate a skill
	All tasks are slightly difficult; it's difficult to guess the answers, and it's not hard to fail to demonstrate a skill when it is called for
	Easy to demonstrate the low level skills, Difficult to demonstrate the higher level skills
	Well-designed tasks to elicit the demonstration of low-level (more concrete) skills, difficult to demonstrate higher level skills
	Easy to demonstrate lower-level skills and tasks for higher-level skills are well-designed

-	*Description*
	Tasks are highly constrained, providing learners with few opportunities to demonstrate abilities not critical to completion of the task; probability of a novice failing to demonstrate a skill when it's not critical to the completion of the task is set to 1
	Tasks provide learners with moderate opportunities to demonstrate abilities not critical to task completion; probability of a novice failing to demonstrate a skill when it's not critical to the completion of the task is set to 1
	Tasks are open-ended, providing learners with ample opportunities to demonstrate abilities not critical to task completion; probability of a novice failing to demonstrate a skill when it's not critical to the completion of the task is set to 1

Sample Pattern



Sample Summary Table

	L1_1	L1_2	L1_3	L2_1	L2_2	L2_3	L3_1	L3_2	L3_3	CL1_1	CL1_2	CL1_3	CL2_1	CL2_2	CL2_3
L1_1	1	1	1	1	0.82	0.28	0.06	0.07	0.07	1	1	0.82	0.29	0.08	0.02
L1_2	1	1	1	1	0.98	0.75	0.2	0.09	0.07	1	0.56	0.34	0.71	0.08	0.03
L1_3	1	1	1	1	1	1	0.49	0.32	0.09	0.43	0.26	0.23	1	0.11	0.03
L2_1	1	1	1	1	1	1	0.94	0.32	0.23	0.45	0.16	0.1	1	0.2	0.09
L2_2	0.82	0.98	1	1	1	1	1	1	0.98	0.09	0.11	0.1	1	0.24	0.08
L2_3	0.28	0.75	1	1	1	1	1	1	1	0.1	0.11	0.1	1	0.78	0.09
L3_1	0.06	0.2	0.49	0.94	1	1	1	1	1	0.03	0.03	0.02	1	0.93	0.2
L3_2	0.07	0.09	0.32	0.32	1	1	1	1	1	0.05	0.05	0.03	1	1	0.22
L3_3	0.07	0.07	0.09	0.23	0.98	1	1	1	1	0.03	0.03	0.02	0.98	1	0.51
CL1_1	1	1	0.43	0.45	0.09	0.1	0.03	0.05	0.03	1	1	1	0.15	0.06	0.01
CL1_2	1	0.56	0.26	0.16	0.11	0.11	0.03	0.05	0.03	1	1	1	0.14	0.06	0.02
CL1_3	0.82	0.34	0.23	0.1	0.1	0.1	0.02	0.03	0.02	1	1	1	0.13	0.05	0.01
CL2_1	0.29	0.71	1	1	1	1	1	1	0.98	0.15	0.14	0.13	1	0.3	0.05
CL2_2	0.08	0.08	0.11	0.2	0.24	0.78	0.93	1	1	0.06	0.06	0.05	0.3	1	0.98
CL2_3	0.02	0.03	0.03	0.09	0.08	0.09	0.2	0.22	0.51	0.01	0.02	0.01	0.05	0.98	1
CT_1	0.94	0.44	0.26	0.11	0.09	0.1	0.03	0.05	0.02	1	1	1	0.11	0.06	0.01
CT_2	1	0.75	0.49	0.29	0.2	0.17	0.06	0.09	0.07	1	1	1	0.2	0.08	0.03
CT_3	1	1	0.94	1	0.44	0.3	0.09	0.14	0.1	1	1	0.99	0.31	0.11	0.07
CT_4	1	1	1	1	1	0.6	0.24	0.23	0.2	0.67	0.46	0.13	0.6	0.18	0.08
CT_5	0.82	0.87	1	1	1	1	1	0.84	0.6	0.24	0.07	0.07	1	0.31	0.13
CT_6	0.08	0.55	0.67	1	1	1	1	1	1	0.05	0.05	0.02	1	0.57	0.2
CT_7	0.06	0.05	0.29	0.18	0.95	1	1	1	1	0.03	0.03	0.01	0.78	1	0.29
CT_8	0.03	0.03	0.05	0.11	0.11	0.54	1	1	1	0.01	0.02	0.01	0.43	1	1
CT_9	0.01	0.02	0.02	0.05	0.05	0.05	0.11	0.13	0.82	0.01	0.01	0.01	0.02	0.8	1
CT_10	0.01	0.01	0.01	0.01	0.02	0.03	0.07	0.07	0.07	0.01	0.01	0.01	0.01	0.07	1
CT_11	0	0	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0	0.01	0	0.01	0.01	0.07
REF	0.06	0.23	0.57	0.71	0.94	0.94	0.45	0.39	0.22	0.02	0.02	0.02	0.92	0.08	0.01

Latent-variable Models

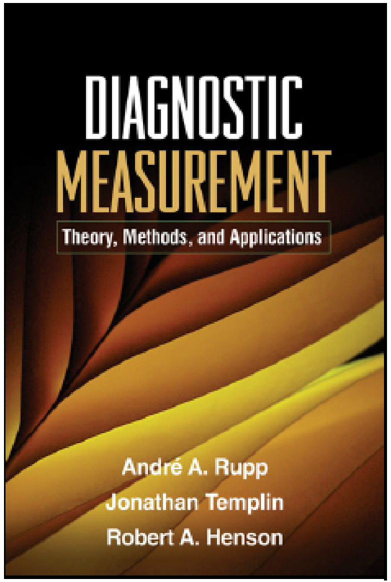
Properties of Latent-variable Models

Selected Advantages

- Provide proper statistical tests for model parameters when assumptions hold
- Correct for measurement error of the epistemic frame elements
- Allow for the incorporation of covariates at different levels
- Allow for the estimation of multi-group and mixture structures

Selected Disadvantages

- Require large player sample sizes relative to non-parametric methods
- Require large number of indicators for each latent variable for reliable estimation
- Require parallel game segment design for appropriate vertical scaling



January 2010
7" x 10" Paperback, 396 Pages
ISBN 978-1-60623-527-0
Cat. #2E3527, $49.00
SPECIAL DISCOUNT PRICE: $41.65

Also Available in Hardcover
ISBN 978-1-60623-528-7
Cat. #2E3528, $72.00

This Title is Part of the **Methodology in the Social Sciences Series**, edited by Todd D. Little.

Diagnostic Measurement

Theory, Methods, and Applications

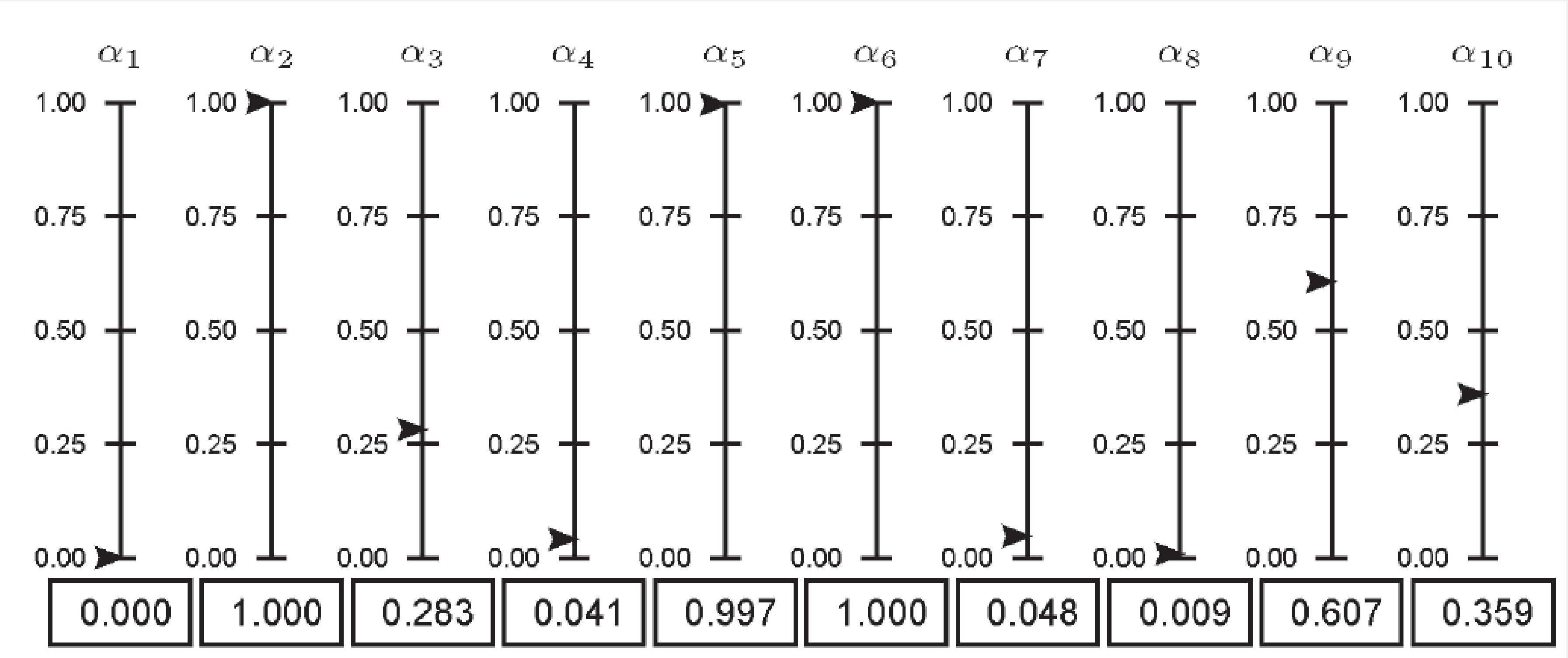
André A. Rupp, Department of Measurement, Statistics, and Evaluation, University of Maryland
Jonathan Templin, Department of Educational Psychology and Instructional Technology, University of Georgia
Robert A. Henson, Department of Educational Research Methodology, The University of North Carolina at Greensboro

"The most authoritative, comprehensive source to date on every important aspect of diagnostic measurement, including theory, methods, and applications….The writing is clear and smooth, making this complex subject matter much more accessible and less intimidating than one might expect."
—**Lihshing Leigh Wang**, School of Education, University of Cincinnati

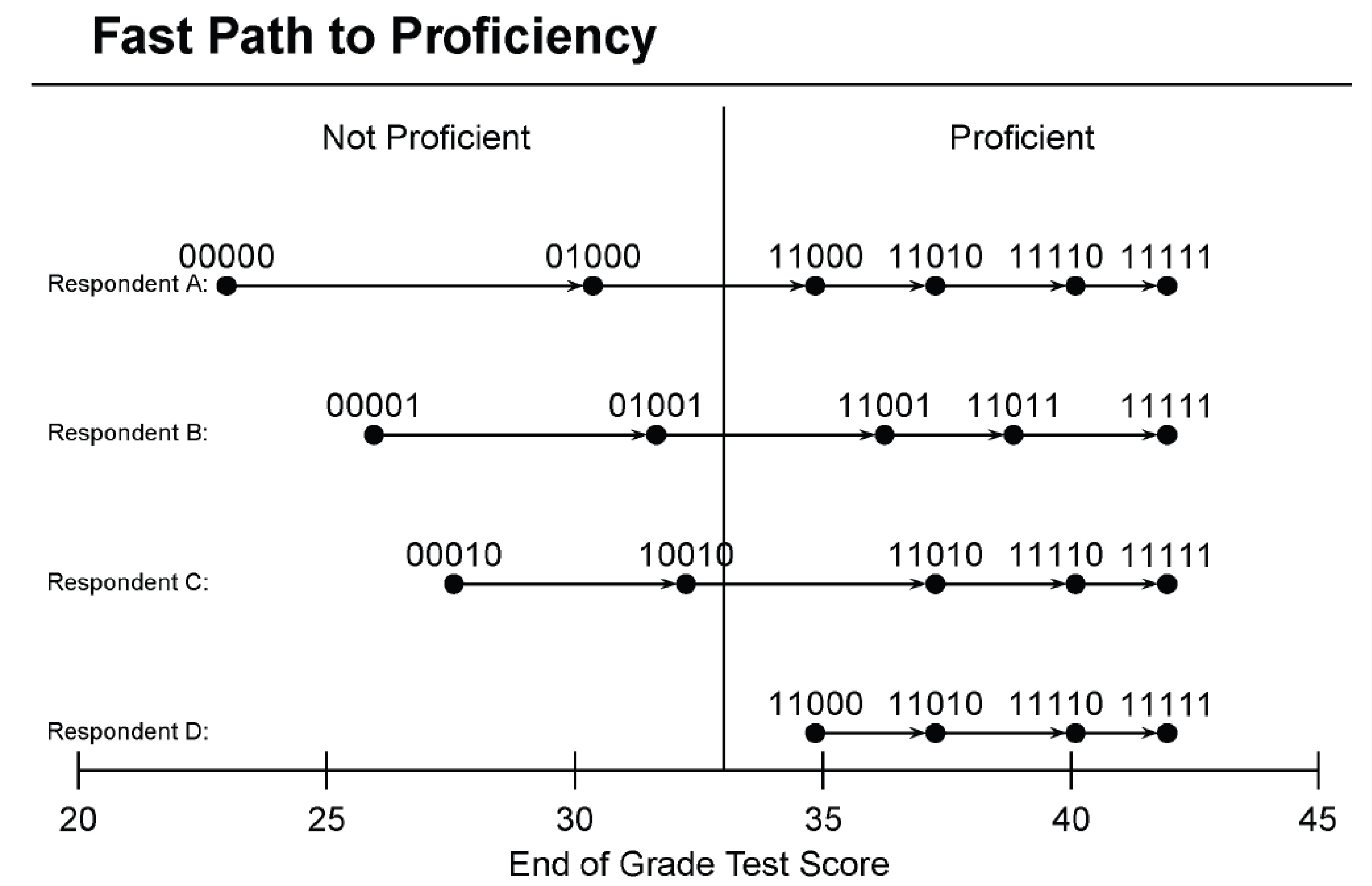
"A real strength of this book is its breadth of coverage. It addresses the importance of embedding diagnostic assessments in a long-term diagnostic process and describes the theoretical underpinnings of diagnostic classification models (DCMs)."
—**Lou DiBello**, Learning Sciences Research Institute, University of Illinois–Chicago

This book provides a comprehensive introduction to the theory and practice of diagnostic classification models (DCMs), which are useful for statistically driven diagnostic decision making. DCMs can be employed in a wide range of disciplines, including educational assessment and clinical psychology. For the first time in a single volume, the authors present the key conceptual underpinnings and methodological foundations for applying these models in practice. Specifically, they discuss a unified approach to DCMs, the mathematical structure of DCMs and their relationship to other latent variable models, and the implementation and estimation of DCMs using *Mplus*. The book's highly accessible language, real-world applications, numerous examples, and clearly annotated equations will encourage professionals and students to explore the utility and statistical properties of DCMs in their own projects.

Sample Output from DCMs



Sample Progression



Example of Mastery Conceptualization

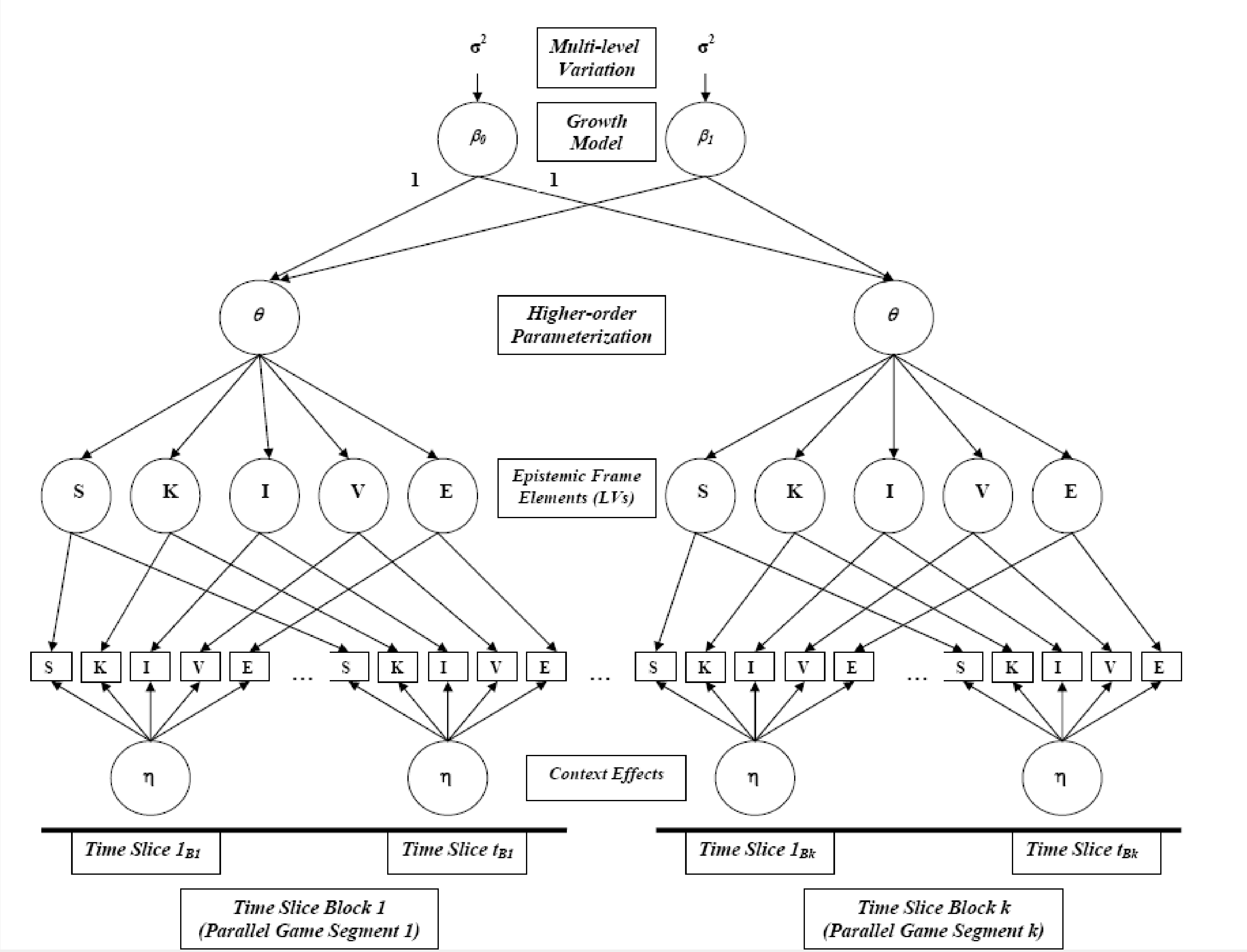
Time Slice	S	K	I	V	E
1	1	1	0	0	0
2	0	1	0	0	0
3	0	0	0	0	0
4	1	1	1	0	1
5	1	0	0	0	0
6	1	1	0	0	0
7	0	0	1	0	0
8	1	1	1	0	1
9	1	1	0	0	0
10	1	1	1	1	0
11	0	0	0	0	0
12	1	1	1	0	1
13	1	0	0	0	0
14	1	1	0	0	0
15	0	0	1	0	0
16	1	1	1	0	1
17	1	0	0	0	0
18	1	1	0	1	1

Game Segment 1

Game Segment 2

Game Segment 3

Longitudinal Latent Variable Models



Further Information

Rupp, A. A., Templin, J., & Henson, R. J. (in press). *Diagnostic measurement: Theory, methods, and applications.* New York: Guilford Press.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Franke, K., Rupp, A. A., & Mislevy, R. J. (in press). *Epistemic network analysis: A prototype for 21st century assessment of learning.* The International Journal of Learning and Media.

Rupp, A. A., Gushta, M., Mislevy, R. J., & Shaffer, D. W. (in press). *Evidence-centered design of epistemic games: Measurement principles for complex learning environments.* Journal of Technology, Learning, and Assessment.

Rupp, A. A., & Templin, J. (2008). *Unique characteristics of cognitive diagnosis models: A comprehensive review of the current state-of-the-art.* Measurement: Interdisciplinary Research & Perspectives, 6, 219-262.

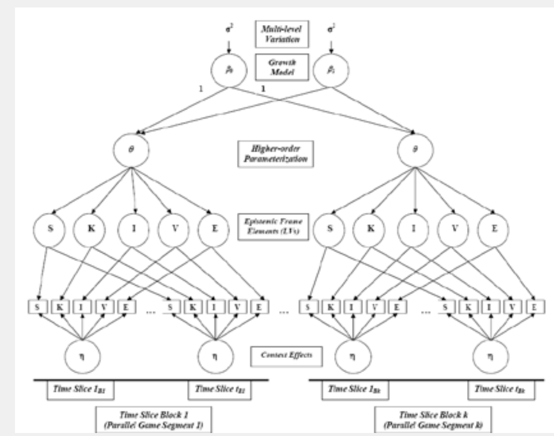
Rupp, A. A., Choi, Y., Gushta, M., Mislevy, R. J., Bagley, E., Nash, P., Hatfield, D., Svarowski, G., & Shaffer, D. (2009). *Modeling learning progressions in epistemic games with epistemic network analysis: Principles for data analysis and generation.* Proceedings from the Learning Progressions in Science Conference held in Iowa City, IA, June 24-26. Available online at http://www.education.uiowa.edu/projects/leaps/

…and David Shaffer's group's website: www.epistemicgames.org

Developing, Modeling, and Representing Domain-specific Expertise via Epistemic Games: A Measurement Person's View

André A. Rupp
EDMS Department
University of Maryland
1230-A Benjamin Building
College Park, MD 20742

E-mail: ruppandr@umd.edu
Phone: (301) 405 – 3623
Fax: (301) 314 – 9245



Physics Education Research Group
University of Maryland, February 22, 2010