Modeling Learning Trajectories with Epistemic Network Analysis: A Simulation-based Investigation of a Novel Analytic Method for Epistemic Games

André A. Rupp, Shauna J. Sweet, & Younyoung Choi
ruppandr@umd.edu
EDMS Department, University of Maryland

Abstract. *Epistemic games* are designed to help players develop domain-specific expertise that characterizes how professionals in a particular domain reason, communicate, and act [1, 11]. To analyze the complex data that arise from these games, a novel analytic method grounded in social network analysis called *epistemic network analysis* (ENA) has been recently proposed [8, 9, 12]. In this paper, we report on preliminary results of a comprehensive simulation study that investigates whether ENA statistics are sensitive to players' differential learning trajectories throughout different game structures under different solution strategies. Preliminary results show a complex emerging picture of the conditions under which one ENA statistic can be suitable for this purpose.

1 Introduction

Learning in the 21st century is increasingly characterized by our ability to make and understand interconnections between concepts, ideas, and conventions across a variety of domains. Consequently, one of the principal challenges of our times is to adequately prepare learners of all ages for challenges in such an increasingly interconnected world, which is heavily permeated by the existence and use of digital tools. Various authors and institutions have proposed taxonomies of so-called *21st-century skills* that are believed to be at the core of the relevant expertise that is required for facing the demands of associated 21st-century tasks [1, 5, 13].

While there is no single definitive list of these skills, most lists focus on expanding traditional concepts of knowledge, skills, and abilities to encompass concepts such as critical and innovative thinking, interpersonal communication and collaboration skills, digital networking and operation skills, intra- and intercultural awareness and identity, and cross-cultural sensibility. Discipline-specific learning as well as learning more generally is not simply restricted to the mastery of concepts and procedures, but includes the ability to think, act, and interact with others in productive ways to solve complex tasks in real-world situations. Becoming an architect, for example, is more than knowing materials properties and tools for computer-aided design. It is being able to see what architects see and being able to frame it in the ways the profession thinks, knowing how to work with and talk with other architects and clients, and using concepts and procedures within the sphere of activities that constitutes architecture. In short, this is what is known as the *epistemic frame* of the discipline [10, 11].

Epistemic games have been developed in recent years to help players develop domain-specific expertise that characterizes how professionals in a particular domain reason, communicate, and act [1, 10]. Although there are many games- and simulation-based opportunities for transforming practices, perceptions, and commitments regarding learning in the 21st century [3, 4], epistemic games are explicitly based on theory of learning in the digital age and are designed to allow learners to develop domain-specific expertise under realistic constraints. For example, learners may learn what it is like to think and act like journalists, artists, business managers, or engineers by using digital learning technologies to solve realistic complex performance tasks. This is accomplished by designing the game in such a way that completing it mimics the core experiences that learners outside the gaming environment would have in a *professional practicum* in the field. The experiences that epistemic games afford and make accessible to learners are characterized by a blend of individual and collaborative work in both real-life and virtual settings.

As one might expect, traditional measurement models with latent variables designed for traditional large-scale assessments struggle to jointly accommodate the complexities of the data that arise from these games. Thus, there are currently no off-the-shelf statistical models that can be applied directly to epistemic games to satisfy the desired scaling and reporting purposes; alternative modeling approaches grounded in non-parametric methods appear to be more promising in this regard.

In this paper, we report on a comprehensive simulation study for investigating one candidate method that has recently been proposed in the literature called *epistemic network analysis* (ENA) [8, 9, 12]. The method is purely descriptive at this point and has been applied to real data collected in several different epistemic games. However, it has not been thoroughly investigated using simulation studies that use conditions representing a wide variety of realistic game-play scenarios. The simulation study reported in this paper is designed to fill this gap to some degree.

1 Basic Principles and Procedures of Epistemic Network Analysis

In epistemic games the sequence of activities in the game is divided into so-called *slices*, *episodes*, or *activity segments*; we will use the term slices in this paper for simplicity. The slices could be defined in a sequential and contiguous fashion based on objective criteria such as time (e.g., 15-minute intervals) or macro-task boundaries (e.g., the beginning and end of a task), but they could also be defined in a sequentially but non-contiguous fashion based on objective criteria such as interactional structures (e.g., pair work versus large-group discussions) or more subjective criteria such as thematic foci (e.g., content analysis of interaction segments).

Once the sequence of activities in the game has been segmented the objective is to extract relevant information about the epistemic frames of the learners from observable products. These products can contain sequences of actions (e.g., sequences of mouse clicks, residing times in segments of the interface), learner products (e.g., report drafts, presentation drafts), or discourse segments (e.g., discussions between pairs of learners, questions posed to mentors). Once the relevant observable products have been identified, the products are coded with respect to relevant evidence about the epistemic frame elements that are activated by the game. The resulting codes constitutes the response data for the game that is used as the input for ENA analyses, which transforms it into numerical and visual representations that can be fed back to the learners and their mentors during the game.

Structure of Resulting Data

The data that arise from the above process can technically be of any nature; previous instantiations of ENA analyses for epistemic games were based on dichotomous (i.e., yes-no / 1-0) codes that indicated whether evidence for a particular epistemic frame element was present. Moreover, the number of epistemic frame elements on which evidence is collected depends on the grain size of the desired feedback; recent instantiations ranged from five macro categories to 19 micro-categories. To succinctly describe the principles and procedures of ENA we will use five macro-categories in this section of the paper, which we will label as *Skills* (S), *Knowledge* (K), *Identity* (I), *Values* (V), and *Epistemology* (E).

Since observable evidence for a learner's reliance on one of the SKIVE elements in his or her epistemic frame is recorded by a '1' for a given slice, a sequence of observations for a single learner throughout the game play may look like the one shown in Table 1.

Table 1. Sequence of observations for single learner

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

For each slice, an *adjacency matrix* is now created, which is a statistical representation of the relational structure between the epistemic frame elements. Adjacency matrices contain entries of '1' whenever two frame elements are used by a learner concurrently within a slice and '0' otherwise; Table 2 shows a sample adjacency matrix for slice 1 for the learner whose data are shown in Table 1.

Table 2. Sample adjacency matrix for a single time slice

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

Note. This adjacency matrix is for slice 1 in Table 1.

Since adjacency matrices are available for each slice, the evidence in them can be accumulated across different slices by simply summing the individual entries in the adjacency matrices across the slices of interest. This resulting matrix is called a *cumulative adjacency matrix*; Table 3 shows such a matrix for the first four slices of the learner from Table 1.

Table 3. Cumulative adjacency matrix

	S	K	I	V	E
S	0	2	1	2	1
K	2	0	1	3	1
I	1	1	0	1	1
V	2	3	1	0	1
E	1	1	1	1	0

This particular process of coding and accumulation implies that the evidence for individual SKIVE elements, in the absence of other SKIVE elements, is discarded. In other words, slices where no SKIVE element is used by a learner carry the same statistical informational value as slices where only one SKIVE element is used in isolation.

ENA Statistics

One can distinguish ENA statistics in terms of whether they provide global marginal information, univariate marginal information, or bivariate information. Due to space limitations for this paper we focus only on one specific global statistic for ease of communication. In terms of notation, we will use $t = 1,\ldots,T$ to refer to slices in reference to the fact that they prototypically denote time or task and because we need the letter s for a task parameter later on, $f = 1,\ldots,F$ to refer to epistemic frame elements, and N to denote the network (i.e., epistemic frame representation) of a particular learner.

Global marginal information is essentially any information that is computed across different rows of the data matrix (i.e., different slices) and different columns of the data matrix (i.e., different SKIVE elements) of a learner. The key global marginal statistics in ENA is the *overall weighted density* (WD) of the network for any given cumulative adjacency matrix, which is computed as follows:

$$WD_t(N) = \sqrt{\tfrac{1}{2}\sum_{f=1}^{F}\sum_{f'=1}^{F'} a_{ff',t}^2}$$

where $a_{ff',t}^2$ is the squared entry in the cumulative adjacency matrix for nodes f and f' at slice t. Note that the total sum is divided by two because the cumulative adjacency matrix is symmetric. The overall weighted network density thus represents the average pair-wise association between nodes in the network that represents the epistemic frame.

In order to gain a better understanding of the statistical properties of ENA statistics, specifically the WD, we designed a simulation study that used principles from latent-variable modeling for data generation. Since ENA is not a parametric model or even a fully formulated non- or semi-parametric model, the purpose of the simulation study was not to assess parameter recovery of this method. Rather, the objective was to generate data according to a plausible mechanism that could separate the influence of task and learner parameters on observed responses and to investigate whether ENA statistics are sensitive to the underlying design characteristics.

3 Methods

In order to simulate data we used principles from modern latent variable models, specifically models in *item response theory* (IRT) [2] and *diagnostic classification models* (DCMs) [6, 7]. In these models, contributions of learner and task characteristics to response probabilities are statistically separated by specifying separable parameters for each.

Specification of Learner Parameters

In this simulation study we chose to frame the data generation from a DCM perspective, which is consistent in principle with the multidimensional IRT perspective. In order to represent different learning trajectories for the different SKIVE nodes we placed different trajectories on the mastery probabilities across the different slices that make up the game play. We distinguished three core sets of growth trends, two of them consisting of linear growth trends and one of them consisting of curvilinear growth trends.

In the first set of linear growth trend, non-zero slopes were modeled with different intercepts corresponding to different initial probabilities of mastery at the beginning of game play. The slopes themselves were determined via linear interpolation by setting the desired final probabilities of mastery at the end of the game play. Crossing three different initial probabilities of (.1, .3, .5) with three different final probabilities of (.6, .8, 1) led to nine different trajectories in this set. In the second set of linear growth trend, zero slopes were modeled that corresponded to trends representing no growth over the course of game play. We used 11 different probabilities corresponding to the set (0, .1, .2,…, 1.0); Figure A1 shows all linear growth trends.

In the set of curvilinear growth trends, slow initial learning followed by quick growth spurts in later stages of the game play and quick learning in early stages of game play followed by a flattened growth trend in later stages of game play were modeled; a total of nine growth trends were modeled. The first subset of growth trends was modeled by using exponential functions with powers 2, 4, and 8, while the second subset of growth trends was modeled by using exponential functions with powers .5, .25, and .125 for a total of six trends; Figure A2 shows the curvilinear growth trends. Overall, we thus modeled a total of 26 growth trends, which provides a broad coverage of different trajectory patterns through the space of mastery probabilities for individual SKIVE nodes.

Specification of Task Parameters

In DCMs the two prototypical task parameters are "slipping" and "guessing" parameters. Slipping parameters, typically denoted by the letter s, represent probabilities of responding inappropriately when learners have mastered a particular required skill or a set of required skills; thus, the reverse probability of $(1 – s)$ is the probability of providing an appropriate response. Guessing parameters, typically denoted by the letter g, represent probabilities of responding appropriately when learners have not mastered a particular required skill or a set of required skills. In other words, $(1 – s)$ and g are the probabilities of responding appropriately when expected and when not as characterized by the mastery status of the learners.

In achievement contexts, the focus is on responding correctly to a particular task, either in absolute terms, when dichotomous scores are used, or in graded terms, when polytomous scores are used. In the contexts of epistemic games this scenario is more complicated because learners are expected to produce efficacious solutions to tasks. That is, learners are expected to produce solutions that are both effective (i.e., they solve the problem at hand) and efficient (i.e., they draw on the key frame elements necessary to solve the problem). Thus, given a particular reference pattern from an expert or mentor for epistemic game play, learners that have high levels of expertise are expected to match the expert's or mentor's pattern as closely as possible. Since experts draw only on epistemic frame elements relevant to a particular task under this conceptualization, the fact that learner does not do so either – and thus produces a '0' in his observed data string for a slice – is not necessarily evidence of a lack of expertise; it may be evidence of emerging expertise. In the context of epistemic games, it is not just '1's in data strings that are desirable but, rather, appropriate sequences of '0's and '1's.

To acknowledge this characteristic of epistemic games we defined four task parameters for this simulation study that are derivatives of such parameters in DCMs. Specifically, we distinguish between guessing and slipping parameters for each SKIVE node and for each of the two possible responses, '0' and '1', separately because both could represent appropriate responses for a given slice. Using α to denote the latent variable that represents a particular epistemic frame element, E to denote the expert response, X to denote the learner response, and a parenthetical superscript to denote whether the slipping or guessing parameter is for the response of '1' or '0' we can define the task parameters as follows:

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

Using these parameter definitions we can now define the marginal probabilities of producing the desired response of '1', $P(X=1|E=1)$, and of '0', $P(X=0|E=0)$, for a particular SKIVE node. It is a function of both mastery states on the node, $\alpha = 1$ (mastery) and $\alpha = 0$, or the respective probabilities of mastery, $P(\alpha = 1)$, and non-mastery, $P(\alpha = 0)$, as well as the task parameters as defined above:

$$\begin{aligned} P(X=1|E=1) &= P(X=1|\alpha=1,E=1)P(\alpha=1)+P(X=1|\alpha=0,E=1)P(\alpha=0) \\ &= 1-s^{(1)}P(\alpha=1)+g^{(1)}P(\alpha=0) \\ P(X=0|E=0) &= P(X=0|\alpha=1,E=0)P(\alpha=1)+P(X=0|\alpha=0,E=0)P(\alpha=0) \\ &= 1-s^{(0)}P(\alpha=1)+g^{(0)}P(\alpha=0) \end{aligned}$$

In this simulation study we set the task parameters deliberately with fixed values to describe an array of different game conditions. Table A1 shows the task parameter values for different conditions as "High" (.25), "Medium" (.15), and "Low" (.05), which covers a range of reasonable values similar to such values in simulation studies for DCMs. We used one reference (i.e., expert) pattern for the initial runs reported here, which came from a real analysis of game data.

Data Generation and Outcome Statistics

The entire simulation study consists of more than 1,000 conditions. All data generation, compiling, and aggregation was done in *R* (www.project-R.org); the *R* code is available from the first author upon request. In the first step we generated the strings of '0's and '1's by constructing the probability matrices using equations 6 and 7, then drawing from a Bernoulli distribution for each cell in the data table, and then matching the outcome of the draw with the expert response. For example, if the expert entry for node '*I*' was '0', the probability for this entry was .86, and a '1' was drawn from the Bernoulli distribution representing the event of interest, a '0' was effectively recoded in the data set as that was the outcome of interest. We replicated this process 100 times.

In the second step we computed what we would call 'empirical confidence bands' for all outcome statistics. For each time slice we computed the lower and upper percentile for the distribution of the ENA statistic of interest across all 100 replications. In this study we used the 2.5th and 97.5th percentiles of the distribution to generate 95% empirical confidence bands. In the third step we then computed the percentage overlap between the confidence bands for learners with different trajectories by counting the number of time slices for which the respective 95% confidence bands overlapped. For example, if the confidence bands overlapped in 62 out of the 87 time slices, this number would have been 71.26%; in general one would expect this overlap to be higher for learners with similar underlying trajectories and lower for learners with different underlying trajectories mediated, of course, by the relative magnitudes of the task parameters for the condition of interest.

4 Preliminary Results

The parameters describing *task difficulty*, $1 - s^{(1)}$ and $g^{(1)}$, are equally low across all tasks indicating that all tasks are well designed. In this context this means that a learner who has mastered a particular SKIVE element will likely apply it when the solution strategy requires it but will not be able to do so when he or she has not mastered the element. At the same time, the primary parameter describing the *task specificity*, $1 - s^{(0)}$, varies across the three conditions reported in this paper while the secondary parameter $g^{(0)}$ is set to 1 across all conditions. The latter specification is necessary because it is unrealistic to allow learners to demonstrate skills when they are not required and they have not mastered them.

In the first set of game conditions (A1) tasks are highly specific. Thus, a learner who has mastered a particular SKIVE element has a low probability of using it unnecessarily when it is not an essential part of the solution strategy for that task. In the second set of conditions (B1) tasks are moderately specific. Thus, a learner who has mastered a particular SKIVE element has a moderately high probability of using it unnecessarily when it is not essential for the solution strategy for that task. In the third set of conditions (C1) tasks are not very specific. Thus, a learner who has mastered a particular SKIVE element has a high probability of using it unnecessarily when it is not essential for the solution strategy for that task. The parameters for tasks are set identically across all time slices for the results presented here but we will also vary them across blocks of tasks in upcoming conditions of the simulation study.

Figures A3, A4, and A5 show the sensitivity of the WD to different learning trajectories under these three separate game conditions. An inspection of the percentage overlap statistic for the 95% confidence bands suggest that the WD can successfully differentiate between learners with different learning trajectories when those trajectories are reasonably distinct in shape. That is, when the underlying learning trajectories are more similar in terms of the probabilities of mastery, the percentage overlap of the confidence bands for the WD is higher and vice versa. This is true within sets of trends (e.g., the percentage overlap of L1_1 with L1_3, L2_3, and L3_3 is 87%, 59%, and 38% respectively) and across sets of trends (e.g., the percentage overlap of L1_1 with CL1_3 and CL2_3 is 59% and 15%, respectively). However, there are also numerous conditions when the learning trajectories are relatively similar in shape and the WD statistic is not able to tell learners apart.

This part of the results would suggest that the WD is a single number that could help differentiate between learners with different underlying trajectories when these are progressing in relatively distinct ways. There is an additional caveat to this statement, however, which can be interpreted both positively and negatively depending on the objective of an analysis. A further inspection of Figures A3 – A5 shows that the relative magnitude of the percentage overlap for any given pair of trajectories is ordered such that it is higher when the tasks are highly specific and lower when the tasks are highly unconstrained. In other words, the statistic becomes more sensitive to trajectory differences when students are able to use SKIVE elements that are not essential for the solution strategy indicated by the expert. This can be viewed as positive if the use of many SKIVE elements are generally encouraged (i.e., when effective yet relatively inefficient solutions are encouraged). It can also be viewed as negative if efficacious solutions are the target of the game play.

5 A Brief Outlook onto Future Work

Due to the space conditions we were only able to present a small snapshot of the results that we already have available and the results that will be available come the conference in June 2010. By that time, different statistics, reference patterns, and all task conditions will have been evaluated. Furthermore, learning trajectories for different SKIVE elements will have been varied, parameters will have been varied across blocks of time slices within a game, and alternative means for data generation will have been explored.

References

[1] Bagley, E., & Shaffer, D. W. (2009). When people get in the way: Promoting civic thinking through epistemic gameplay. *International Journal of Gaming and Computer-mediated Simulations, 1,* 36-52.

[2] de Ayala, R. J. (2009). *The theory and practice of item response theory.* New York: The Guilford Press.

[3] Gee, J. P. (2003). *What video games have to teach us about learning and literacy.* New York: Palgrave / Macmillan.

[4] Gibson, D., Aldrich, C., & Prensky, M. (Eds.). (2006). *Games and simulations in online learning: Research and development frameworks.* Hershey, PA: Information Science Publishing.

[5] Partnership for 21st Century Skills (2008). *21st century skills, education, and competitiveness: A resource and policy guide.* Tuscon, AZ. Available online at www.21stcenturykills.org

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

[7] Rupp, A. A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York: Guilford Press.

[8] Rupp, A. A., Gushta, M., Mislevy, R. J., & Shaffer, D. W. (2010). Evidence-centered design of epistemic games: Measurement principles for complex learning environments. *Journal of Technology, Learning, and Assessment*, *8*(4). Available online at http://escholarship.bc.edu/jtla/vol8/4

[9] 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.

[10] Shaffer, D. W. (2006a). *How computer games help children learn.* New York: Palgrave / Macmillan.

[11] Shaffer, D. W. (2006b). Epistemic frames for epistemic games. *Computers and Education*. 46(3), 223-234.

[12] Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Franke, K., Rupp, A. A., & Mislevy, R. J. (2009). Epistemic network analysis: A prototype for 21st century assessment of learning. *International Journal of Learning Media, 1(2)*, 33-53.

[13] Shute, V. J., Dennen, V. P., Kim, Y.-J., Donmez, O., & Wang, C.-Y. (in press). 21st century assessment to promote 21st century learning: The benefits of blinking. In J. Gee (Ed.), *Games, learning, assessment.* Boston, MA: MIT Press.

Author Note

This work was made possible, in part, by a grant from the Support Program for Advancing Research and Collaboration (SPARC) within the College of Education at the University of Maryland as well as, in part, by a grant from the John D. And Catherine T. MacArthur foundation awarded to Arizona State University (07-90185-000-HCD) and two grants from the National Science Foundation awarded to the University of Wisconsin at Madison (DRL-0918409 and DRL-0946372). The opinions, findings, and conclusions or recommendations in this paper are those of the authors and do not necessarily reflect the views of the funding agencies, cooperating institutions, or other individuals.

Appendix

Table A1. Task parameters of SKIVE nodes across difficulty and specificity conditions

Difficulty 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 (i.e., all tasks are well designed for slipping and guessing across skills)
2	High	High	High	High	High	High	High	High	High	High	All tasks are poorly designed (i.e., all tasks are poorly designed for slipping and guessing across skills)
3	Low	High	Low	High	Low	High	Low	High	Low	High	All tasks are moderately well designed (i.e., all tasks are well designed for slipping and poorly designed for guessing across skills)
4	High	Low	High	Low	High	Low	High	Low	High	Low	All tasks are moderately well designed (i.e., all tasks are poorly designed for slipping and well designed for guessing across skills)
5	Low	High	Low	High	High	Low	High	Low	High	Low	Tasks are differentially well designed (i.e., tasks are poorly designed for guessing but well designed for slipping on basic skills but well designed for guessing and poorly designed for slipping for complex skills)
6	Low	Low	Low	Low	High	Low	High	Low	High	Low	These tasks are differentially well designed (i.e., tasks are well designed for both guessing and slipping for basic skills while they are well designed for guessing but poorly designed for slipping for complex skills)
7	Low	High	Low	High	Low	Low	Low	Low	Low	Low	These tasks are differentially well designed (i.e., tasks are well designed for slipping but poorly designed for guessing for basic skills but well designed for both guessing and slipping for complex skills)

(*continued*)

Specificity	S		K		I		V		E		Description
Condition	$s^{(0)}$	$g^{(0)}$	$s^{(0)}$	$g^{(0)}$	$s^{(0)}$	$g^{(0)}$	$s^{(0)}$	$g^{(0)}$	$s^{(0)}$	$g^{(0)}$	
A	Low	1	Low	1	Low	1	Low	1	Low	1	All tasks are highly specific (i.e., they provide few opportunities to demonstrate skills that have not been mastered or make successful suppression of skills for an efficient response easy)
B	Medium	1	Medium	1	Medium	1	Medium	1	Medium	1	All tasks are moderately specific (i.e., they provide some opportunities to demonstrate skills that have not been mastered or make successful suppression of skills for an efficient response moderately difficult)
C	High	1	High	1	High	1	High	1	High	1	All tasks are not very specific (i.e., they provide many opportunities to demonstrate skills that have not been mastered or make successful suppression of skills for an efficient response very difficult)

Note: High = .25, Medium = .15, Low = .05. Conditions A1, B1, and C1 are created by combining the difficulty condition 1 with the specificity conditions A, B, and C.

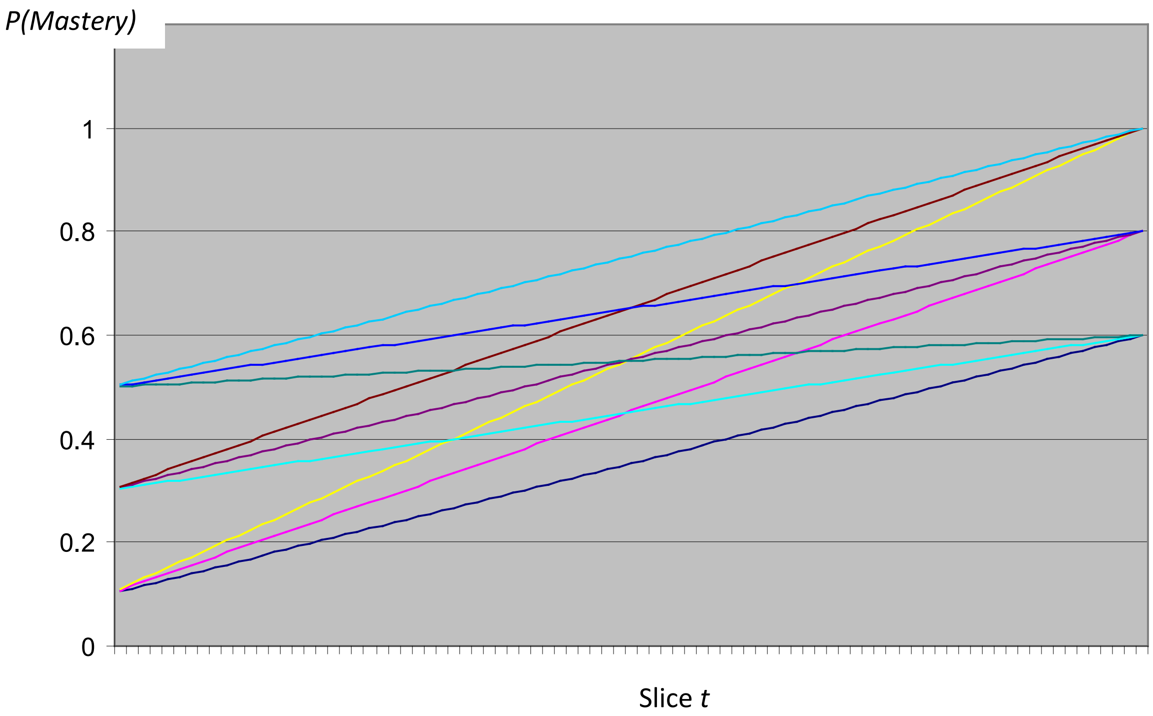


Figure A1. Linear growth trends for mastery probabilities of learners for individual nodes.

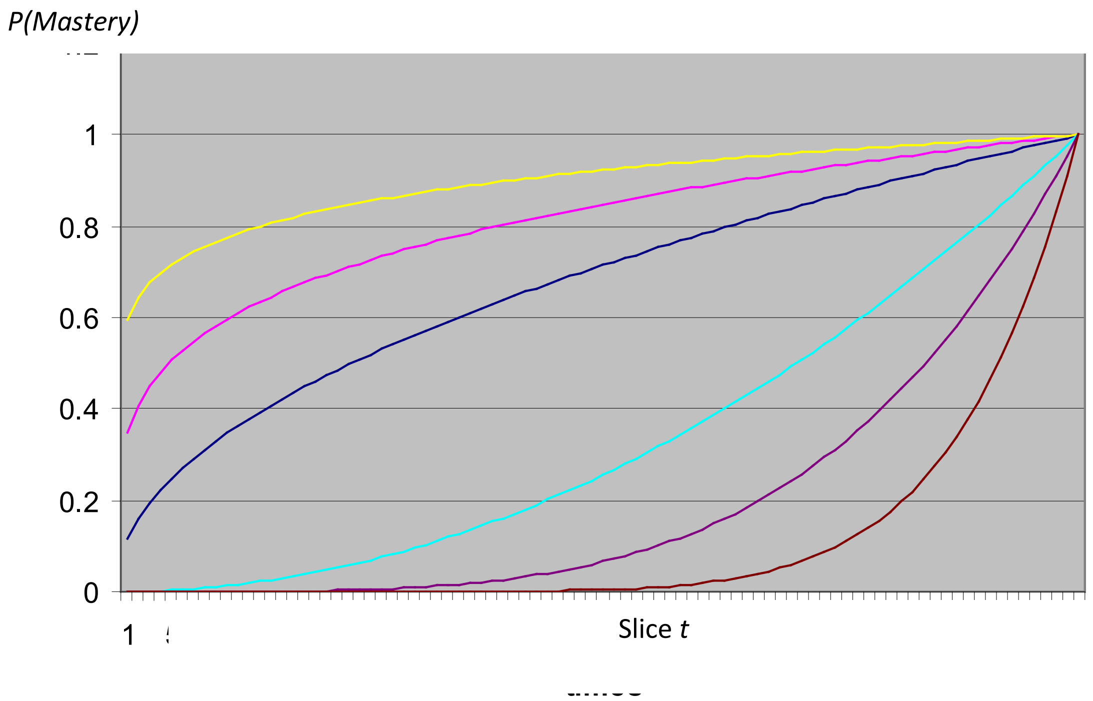
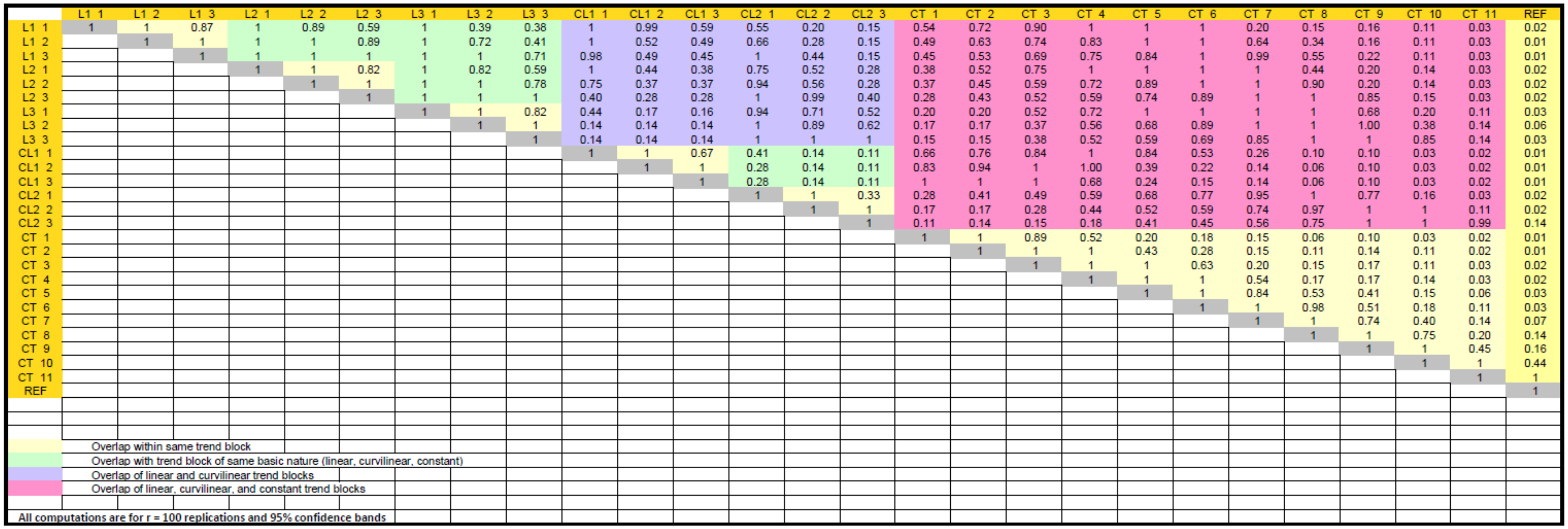


Figure A2. Curvilinear growth trends for mastery probabilities of learners for individual nodes.

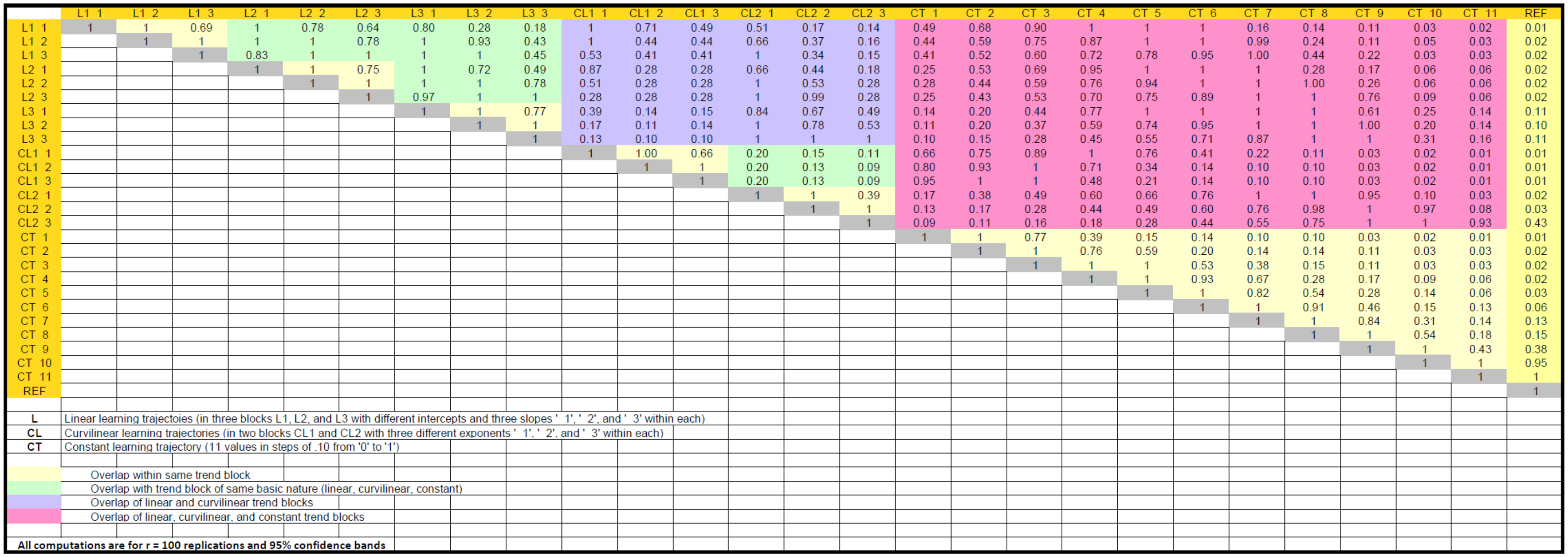


	L1 1	L1 2	L1 3	L2 1	L2 2	L2 3	L3 1	L3 2	L3 3	CL1 1	CL1 2	CL1 3	CL2 1	CL2 2	CL2 3	CT 1	CT 2	CT 3	CT 4	CT 5	CT 6	CT 7	CT 8	CT 9	CT 10	CT 11	REF
L1 1	1	1	0.87	1	0.89	0.59	1	0.39	0.38	1	0.99	0.59	0.55	0.20	0.15	0.54	0.72	0.90	1	1	1	0.20	0.15	0.16	0.11	0.03	0.02
L1 2		1	1	1	1	0.89	1	0.72	0.41	1	0.52	0.49	0.66	0.28	0.15	0.49	0.63	0.74	0.83	1	1	0.64	0.34	0.16	0.11	0.03	0.01
L1 3			1	1	1	1	1	1	0.71	0.98	0.49	0.45	1	0.44	0.15	0.45	0.53	0.69	0.75	0.84	1	0.99	0.55	0.22	0.11	0.03	0.01
L2 1				1	1	0.82	1	0.82	0.59	1	0.44	0.38	0.75	0.52	0.28	0.38	0.52	0.75	1	1	1	1	0.44	0.20	0.14	0.03	0.02
L2 2					1	1	1	1	0.78	0.75	0.37	0.37	0.94	0.56	0.28	0.37	0.45	0.59	0.72	0.89	1	1	0.90	0.20	0.14	0.03	0.02
L2 3						1	1	1	1	0.40	0.28	0.28	1	0.99	0.40	0.28	0.43	0.52	0.59	0.74	0.89	1	1	0.85	0.15	0.03	0.02
L3 1							1	1	0.82	0.44	0.17	0.16	0.94	0.71	0.52	0.20	0.20	0.52	0.72	1	1	1	1	0.68	0.20	0.11	0.03
L3 2								1	1	0.14	0.14	0.14	1	0.89	0.62	0.17	0.17	0.37	0.56	0.68	0.89	1	1	1.00	0.38	0.14	0.06
L3 3									1	0.14	0.14	0.14	1	1	1	0.15	0.15	0.38	0.52	0.59	0.69	0.85	1	1	0.85	0.14	0.03
CL1 1										1	1	0.67	0.41	0.14	0.11	0.66	0.76	0.84	1	0.84	0.53	0.26	0.10	0.10	0.03	0.02	0.01
CL1 2											1	1	0.28	0.14	0.11	0.83	0.94	1	1.00	0.39	0.22	0.14	0.06	0.10	0.03	0.02	0.01
CL1 3												1	0.28	0.14	0.11	1	1	1	0.68	0.24	0.15	0.14	0.06	0.10	0.03	0.02	0.01
CL2 1													1	1	0.33	0.28	0.41	0.49	0.59	0.68	0.77	0.95	1	0.77	0.16	0.03	0.02
CL2 2														1	1	0.17	0.17	0.28	0.44	0.52	0.59	0.74	0.97	1	1	0.11	0.02
CL2 3															1	0.11	0.14	0.15	0.18	0.41	0.45	0.56	0.75	1	1	0.99	0.14
CT 1																1	1	0.89	0.52	0.20	0.18	0.15	0.06	0.10	0.03	0.02	0.01
CT 2																	1	1	1	0.43	0.28	0.15	0.11	0.14	0.11	0.02	0.01
CT 3																		1	1	1	0.63	0.20	0.15	0.17	0.11	0.03	0.02
CT 4																			1	1	1	0.54	0.17	0.17	0.14	0.03	0.02
CT 5																				1	1	0.84	0.53	0.41	0.15	0.06	0.03
CT 6																					1	1	0.98	0.51	0.18	0.11	0.03
CT 7																						1	1	0.74	0.40	0.14	0.07
CT 8																							1	1	0.75	0.20	0.14
CT 9																								1	1	0.45	0.16
CT 10																									1	1	0.44
CT 11																										1	1
REF																											1

🟨	Overlap within same trend block
🟩	Overlap with trend block of same basic nature (linear, curvilinear, constant)
🟪	Overlap of linear and curvilinear trend blocks
🟥	Overlap of linear, curvilinear, and constant trend blocks

All computations are for r = 100 replications and 95% confidence bands

Figure A3. Percent overlap of 95% confidence bands for WD for different learning trajectories under game condition A1.



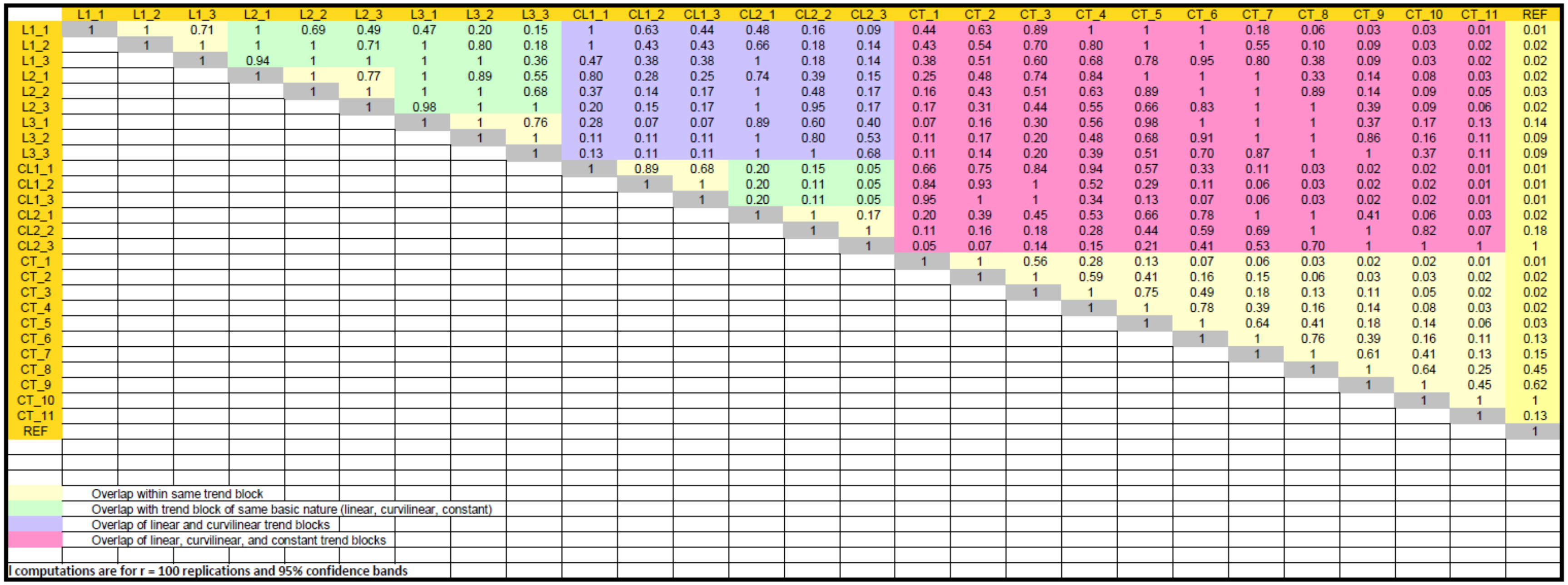
	L1 1	L1 2	L1 3	L2 1	L2 2	L2 3	L3 1	L3 2	L3 3	CL1 1	CL1 2	CL1 3	CL2 1	CL2 2	CL2 3	CT 1	CT 2	CT 3	CT 4	CT 5	CT 6	CT 7	CT 8	CT 9	CT 10	CT 11	REF
L1 1	1	1	0.69	1	0.78	0.64	0.80	0.28	0.18	1	0.71	0.49	0.51	0.17	0.14	0.49	0.68	0.90	1	1	1	0.16	0.14	0.11	0.03	0.02	0.01
L1 2		1	1	1	1	0.78	1	0.93	0.43	1	0.44	0.44	0.66	0.37	0.16	0.44	0.59	0.75	0.87	1	1	0.99	0.24	0.11	0.05	0.03	0.02
L1 3			1	0.83	1	1	1	1	0.45	0.53	0.41	0.41	1	0.34	0.15	0.41	0.52	0.60	0.72	0.78	0.95	1.00	0.44	0.22	0.03	0.03	0.02
L2 1				1	1	0.75	1	0.72	0.49	0.87	0.28	0.28	0.66	0.44	0.18	0.25	0.53	0.69	0.95	1	1	1	0.28	0.17	0.06	0.06	0.02
L2 2					1	1	1	1	0.78	0.51	0.28	0.28	1	0.53	0.28	0.28	0.44	0.59	0.76	0.94	1	1	1.00	0.26	0.06	0.06	0.02
L2 3						1	0.97	1	1	0.28	0.28	0.28	1	0.99	0.28	0.25	0.43	0.53	0.70	0.75	0.89	1	1	0.76	0.09	0.06	0.02
L3 1							1	1	0.77	0.39	0.14	0.15	0.84	0.67	0.49	0.14	0.20	0.44	0.77	1	1	1	1	0.61	0.25	0.14	0.11
L3 2								1	1	0.17	0.11	0.14	1	0.78	0.53	0.11	0.20	0.37	0.59	0.74	0.95	1	1	1.00	0.20	0.14	0.10
L3 3									1	0.13	0.10	0.10	1	1	1	0.10	0.15	0.28	0.45	0.55	0.71	0.87	1	1	0.31	0.16	0.11
CL1 1										1	1.00	0.66	0.20	0.15	0.11	0.66	0.75	0.89	1	0.76	0.41	0.22	0.11	0.03	0.02	0.01	0.01
CL1 2											1	1	0.20	0.13	0.09	0.80	0.93	1	0.71	0.34	0.14	0.10	0.10	0.03	0.02	0.01	0.01
CL1 3												1	0.20	0.13	0.09	0.95	1	1	0.48	0.21	0.14	0.10	0.10	0.03	0.02	0.01	0.01
CL2 1													1	1	0.39	0.17	0.38	0.49	0.60	0.66	0.76	1	1	0.95	0.10	0.03	0.02
CL2 2														1	1	0.13	0.17	0.28	0.44	0.49	0.60	0.76	0.98	1	0.97	0.08	0.03
CL2 3															1	0.09	0.11	0.16	0.18	0.28	0.44	0.55	0.75	1	1	0.93	0.43
CT 1																1	1	0.77	0.39	0.15	0.14	0.10	0.10	0.03	0.02	0.01	0.01
CT 2																	1	1	0.76	0.59	0.20	0.14	0.14	0.11	0.03	0.03	0.02
CT 3																		1	1	1	0.53	0.38	0.15	0.11	0.03	0.03	0.02
CT 4																			1	1	0.93	0.67	0.28	0.17	0.09	0.06	0.02
CT 5																				1	1	0.82	0.54	0.28	0.14	0.06	0.03
CT 6																					1	1	0.91	0.46	0.15	0.13	0.06
CT 7																						1	1	0.84	0.31	0.14	0.13
CT 8																							1	1	0.54	0.18	0.15
CT 9																								1	1	0.43	0.38
CT 10																									1	1	0.95
CT 11																										1	1
REF																											1

L Linear learning trajectoies (in three blocks L1, L2, and L3 with different intercepts and three slopes ' 1', ' 2', and ' 3' within each)
CL Curvilinear learning trajectories (in two blocks CL1 and CL2 with three different exponents ' 1', ' 2', and ' 3' within each)
CT Constant learning trajectory (11 values in steps of .10 from '0' to '1')

🟨 Overlap within same trend block
🟩 Overlap with trend block of same basic nature (linear, curvilinear, constant)
🟪 Overlap of linear and curvilinear trend blocks
🟥 Overlap of linear, curvilinear, and constant trend blocks

All computations are for r = 100 replications and 95% confidence bands

Figure A4. Percent overlap of 95% confidence bands for WD for different learning trajectories under game condition B1.



	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	CT_1	CT_2	CT_3	CT_4	CT_5	CT_6	CT_7	CT_8	CT_9	CT_10	CT_11	REF
L1_1	1	1	0.71	1	0.69	0.49	0.47	0.20	0.15	1	0.63	0.44	0.48	0.16	0.09	0.44	0.63	0.89	1	1	1	0.18	0.06	0.03	0.03	0.01	0.01
L1_2		1	1	1	1	0.71	1	0.80	0.18	1	0.43	0.43	0.66	0.18	0.14	0.43	0.54	0.70	0.80	1	1	0.55	0.10	0.09	0.03	0.02	0.02
L1_3			1	0.94	1	1	1	1	0.36	0.47	0.38	0.38	1	0.18	0.14	0.38	0.51	0.60	0.68	0.78	0.95	0.80	0.38	0.09	0.03	0.02	0.02
L2_1				1	1	0.77	1	0.89	0.55	0.80	0.28	0.25	0.74	0.39	0.15	0.25	0.48	0.74	0.84	1	1	1	0.33	0.14	0.08	0.03	0.02
L2_2					1	1	1	1	0.68	0.37	0.14	0.17	1	0.48	0.17	0.16	0.43	0.51	0.63	0.89	1	1	0.89	0.14	0.09	0.05	0.03
L2_3						1	0.98	1	1	0.20	0.15	0.17	1	0.95	0.17	0.17	0.31	0.44	0.55	0.66	0.83	1	1	0.39	0.09	0.06	0.02
L3_1							1	1	0.76	0.28	0.07	0.07	0.89	0.60	0.40	0.07	0.16	0.30	0.56	0.98	1	1	1	0.37	0.17	0.13	0.14
L3_2								1	1	0.11	0.11	0.11	1	0.80	0.53	0.11	0.17	0.20	0.48	0.68	0.91	1	1	0.86	0.16	0.11	0.09
L3_3									1	0.13	0.11	0.11	1	1	0.68	0.11	0.14	0.20	0.39	0.51	0.70	0.87	1	1	0.37	0.11	0.09
CL1_1										1	0.89	0.68	0.20	0.15	0.05	0.66	0.75	0.84	0.94	0.57	0.33	0.11	0.03	0.02	0.02	0.01	0.01
CL1_2											1	1	0.20	0.11	0.05	0.84	0.93	1	0.52	0.29	0.11	0.06	0.03	0.02	0.02	0.01	0.01
CL1_3												1	0.20	0.11	0.05	0.95	1	1	0.34	0.13	0.07	0.06	0.03	0.02	0.02	0.01	0.01
CL2_1													1	1	0.17	0.20	0.39	0.45	0.53	0.66	0.78	1	1	0.41	0.06	0.03	0.02
CL2_2														1	1	0.11	0.16	0.18	0.28	0.44	0.59	0.69	1	1	0.82	0.07	0.18
CL2_3															1	0.05	0.07	0.14	0.15	0.21	0.41	0.53	0.70	1	1	1	1
CT_1																1	1	0.56	0.28	0.13	0.07	0.06	0.03	0.02	0.02	0.01	0.01
CT_2																	1	1	0.59	0.41	0.16	0.15	0.06	0.03	0.03	0.02	0.02
CT_3																		1	1	0.75	0.49	0.18	0.13	0.11	0.05	0.02	0.02
CT_4																			1	1	0.78	0.39	0.16	0.14	0.08	0.03	0.02
CT_5																				1	1	0.64	0.41	0.18	0.14	0.06	0.03
CT_6																					1	1	0.76	0.39	0.16	0.11	0.13
CT_7																						1	1	0.61	0.41	0.13	0.15
CT_8																							1	1	0.64	0.25	0.45
CT_9																								1	1	0.45	0.62
CT_10																									1	1	1
CT_11																										1	0.13
REF																											1

🟨 Overlap within same trend block
🟩 Overlap with trend block of same basic nature (linear, curvilinear, constant)
🟪 Overlap of linear and curvilinear trend blocks
🟥 Overlap of linear, curvilinear, and constant trend blocks

l computations are for r = 100 replications and 95% confidence bands

Figure A5. Percent overlap of 95% confidence bands for WD for different learning trajectories under game condition C1.