An Elicitation Tool for Conditional Probability Tables (CPT) for Physics Playground

Russell Almond, Seyfullah Tingir, Xi Lu, Chen Sun, Seyedahmad Rahimi

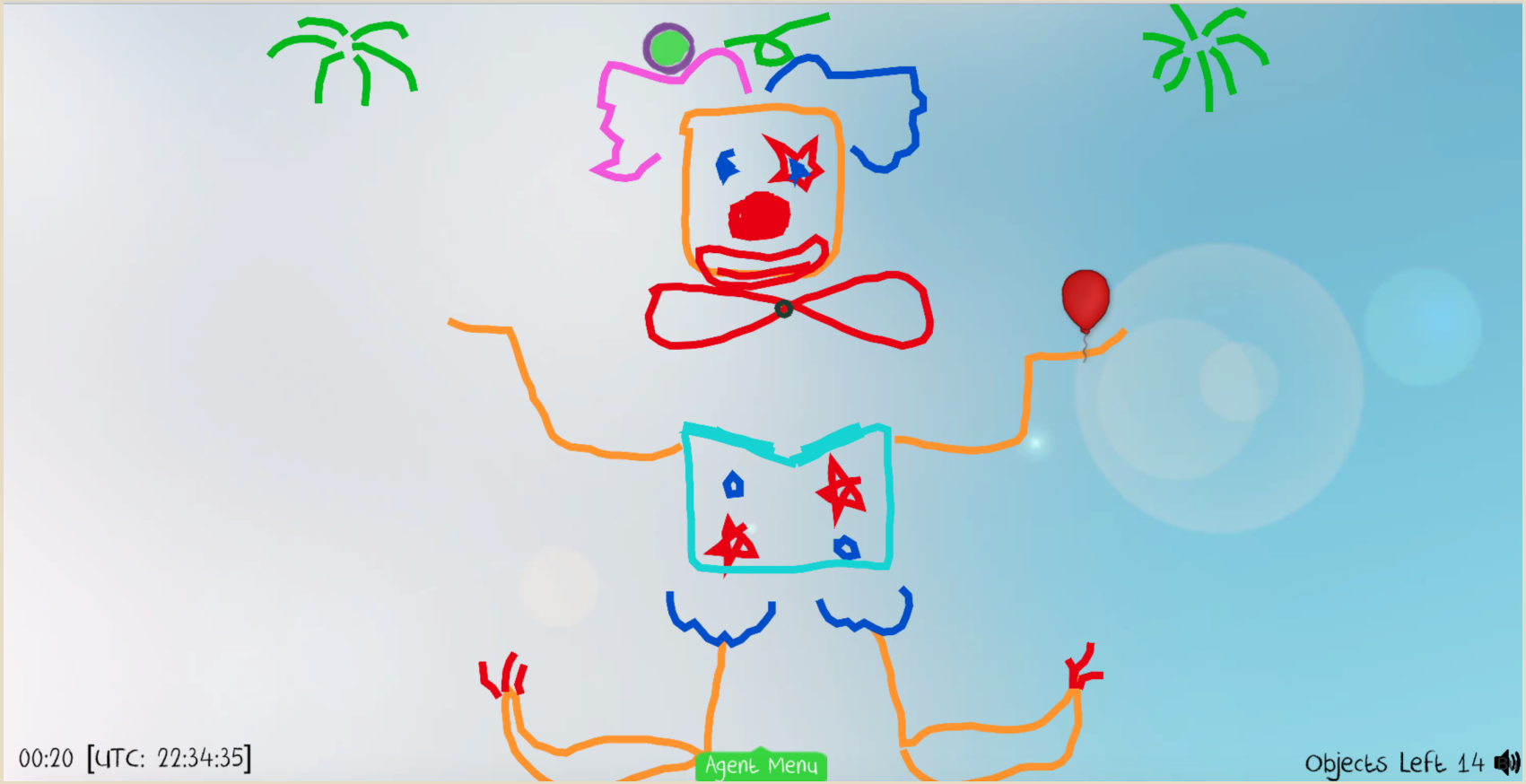
Educational Psychology and Learning Systems
College of Education
Florida State University

Bayesian Modeling Application Workshop 2017

Physics Playground

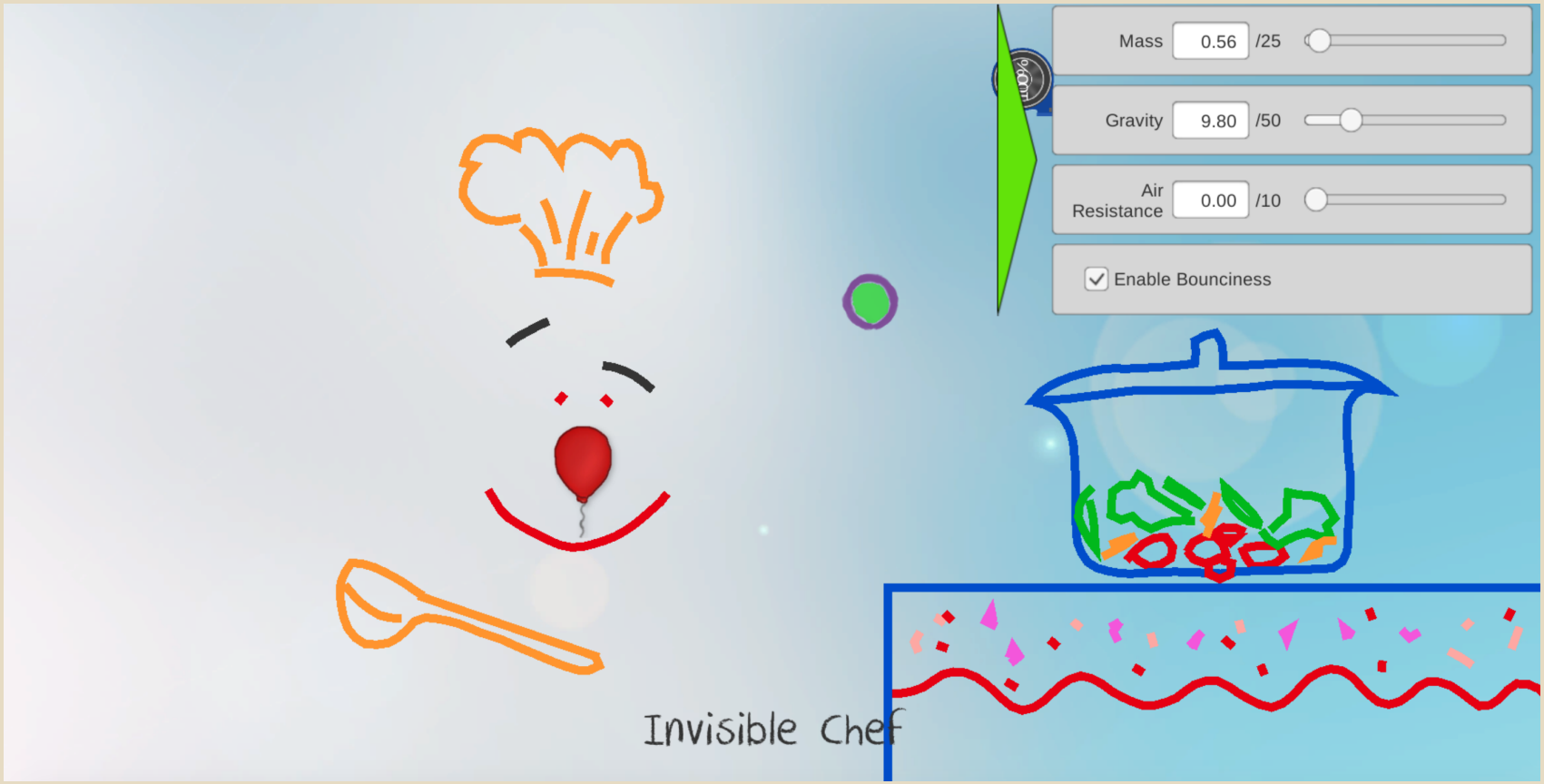
- Students explore a 2D world which follows laws of Physics.
- Can use performance in game to learn about students' knowledge of physics. (Bayes net used for scoring.)
- Version 2.0 adds feedback and instruction, new manipulation levels.
- Eventually, Bayes net will be used for sequencing of levels, triggering feedback.
- Game Demo: `http://www.empiricalgames.org/ppunity5/LE/1.2.8/demo.html`
- Level Editor: `http://www.empiricalgames.org/ppunity5/LE/1.3.1/demo.html`

Drawing Level



Goal: Get the ball to the balloon by drawing objects on the screen.

Manipulation Level



Goal: Get the ball to the balloon by manipulating, the ball's mass, gravity, air resistance and bounciness.

Physics Playground Project Team

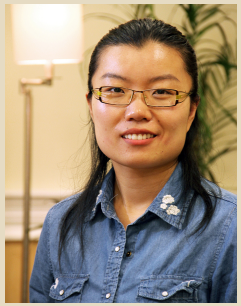
Val Shute (PI)

Fengfeng Ke (co-PI)

Russell Almond (co-PI)

NOT Pictured:
Adam LaMee
Don
Franceschetti

Ahmad Rahimi



Chen Sun

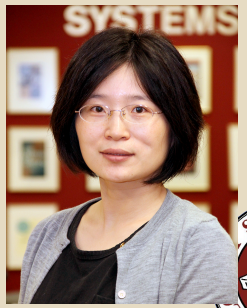
Ginny Smith

Lukas Liu

Seyfullah Tingir

Weinan Zhao

Xinhao Xu



Xi Lu

Hub and Spoke Model

- Educational Models have a special structure.
- Central *Competency Model* or *Proficiency Model* contains variables of interest.
- A collection of *Evidence Models* corresponding to *tasks* or activities that a student does.
  - These would be called *tests* in a medical or system diagnosis application.
- Evidence models link observable variables to variables in the center competency model.
  - Model fragment: competency variables are just stubs (references).
- A test or *assessment* is made up of a collection of tasks.
  - Bayes net for assessment (competency plus relevant evidence models) is called a *motif*.

Competency Model

- Defines scope and grain size of project.
- Co-ordinates work of teams (level developer, learning supports, scoring, physics experts).
- Critical to define both what variables mean and how they relate to the target population.
- *Middle school students have not yet learned vectors!*
- Represented with a Bayesian network.

Steps in developing the Competency Model

- Defining the variables
- Associating competencies with game levels (Q-Matrix)
- Bayesian network structure
- Parameterizing the network
- Translating the Numbers into Natural Language and Validating with Experts
- Market Basket Validation

Defining Latent Variables

- Key variables are latent: cannot be observed directly.
- Definition is implicit through tasks.
- Construct Map (Wilson, 2006). For High, Med and Low levels:
  - What kind of *tasks*?
  - What qualities of *performance*?
  - What *personal* qualities?
- Add Evidence column to provide ideas for how to test.
- Competencies are usually defined hierarchically (Math and Science).
- Grain size is important.
  - 3–6 observations needed for each competency variable.
- Target population is important.
- *Joint effort of Physics Experts and Design Team*

Variable Definition Spreadsheet

Table 1: ES Spreadsheet

Competency	Sub-competency	Description	Evidence
Force and Motion	Newton's 1st Law	Static equilibrium (a=0 and v=0)	Player applies or adjusts a force (e.g., nudge, blow, gravity, air resistance) to keep an object stationary in at least one dimension.
Force and Motion	Newton's 2nd Law	Net force and acceleration are directly related	Player applies or adjusts a force acting on an object to cause it to accelerate at a desired rate.
Linear Momentum	Properties of momentum	Momentum is directly related to mass	Player adjusts the mass an object to affect the amount of momentum it transfers to a second object after the two collide.
Energy	Energy can transfer	Energy can transform from one type to another (e.g., GPE to KE)	Player changes parameters (e.g., mass, position, speed) to transform more or less energy of one type to another (e.g., KE, GPE, EPE) of the same object.
Torque	Properties of torque	Force and torque are directly related.	Player adjusts the magnitude of a force to cause a corresponding change in the magnitude of a torque.
Science and Engineering Practices	Use iterative design to solve a problem	Solve a problem by making variations on previous strategies	Player makes successive adjustments of the same parameter to solve a level.

Q-matrix

- Variables are defined by tasks used to measure them.
- Associate tasks (game levels) with competencies.
- Make sure we have good balance.
- Find out if competencies are really distinguishable.
- $q_{jk} = 1$ iff Competency k is relevant for Task j.
- Column sums provide information about coverage.
- *Initial draft by design team, reviewed by Physics experts.*

Q-matrix excerpt.

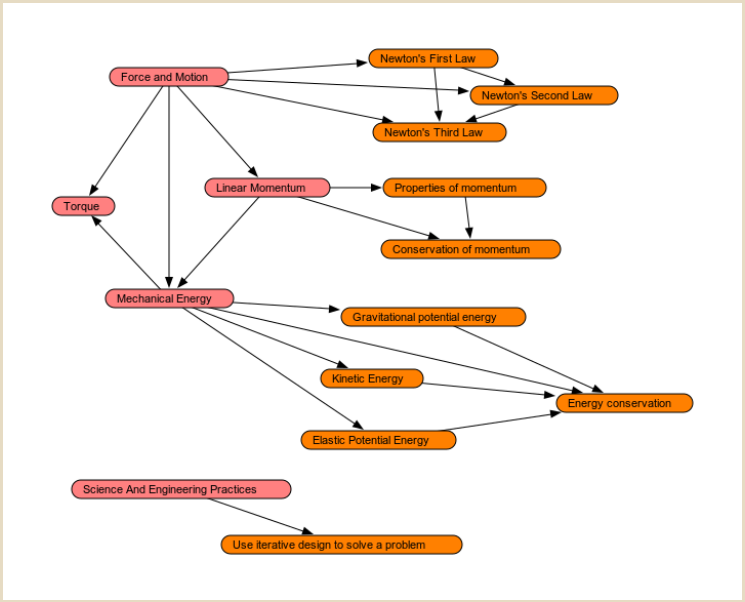
Table 2: Q-matrix designed by the learning support team

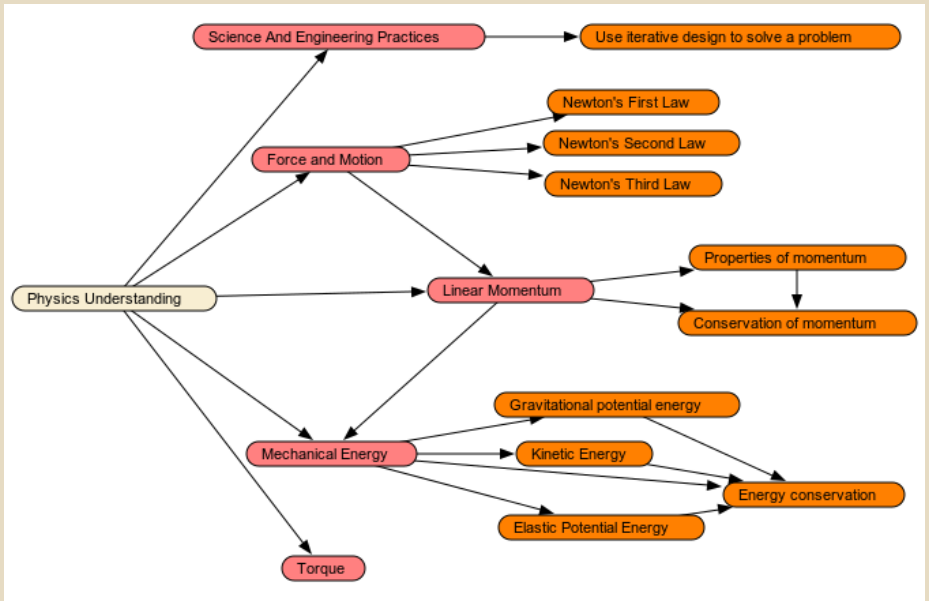
	Force and Motion		
	Newton's 1st Law	Newton's 2nd Law	Newton's 3rd Law
On the Upswing	1	0	0
Lead the Ball	0	0	0
Scale	1	0	0
Spider Web	0	0	0
On the Upswing	0	0	0

Defining the structure

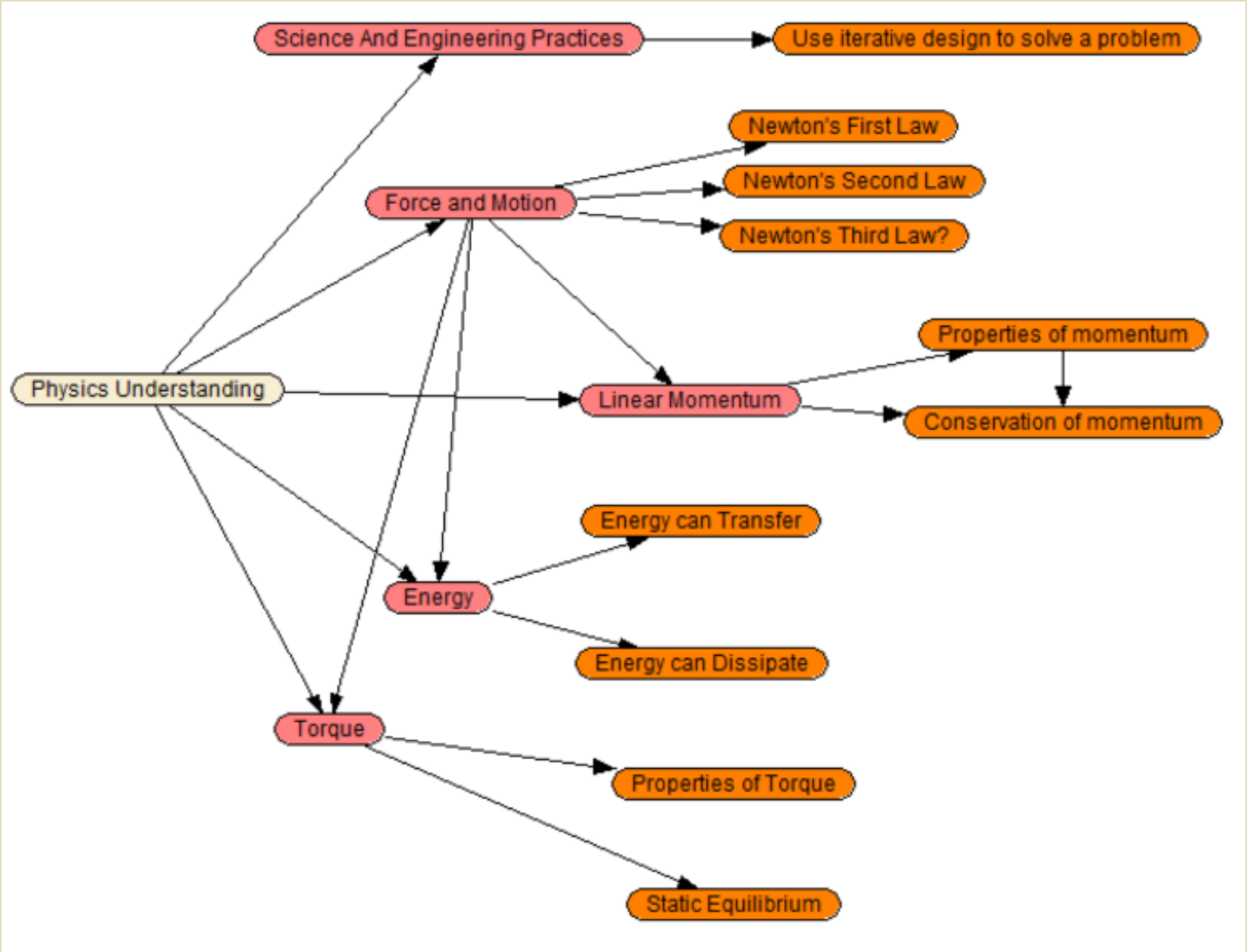
- Working in Netica, but as a drawing tool.
- Experts tend to give hierarchical models, Bayes nets are about conditional independence.
- Modeling team built two candidate networks.
- Presented to Physics and Design teams to have them review.

Two Candidate Structures





Final (Version 6) Structure



DiBello–Normal Models

- Based on suggestion from expert to elicit parameters of a regression model
- Map states of parent variables onto point on a $N(0,1)$ scale.
- Elicit *discrimination* (slope) and *difficulty* (negative intercept) parameters.
- Elicit a residual variance *link scale* parameter, a_0

$$R^2 = \frac{\sum a_k^2/K}{\sum a_k^2/K + a_0^2}$$

- Use these to computer CPT (code in `CPTtools` package).
- `Peanut` package provides support for associating the parameters with Pnode (parameterized nodes) objects.

Getting Expert Validation

- Modeling Team produced first pass numbers based on some general guesses.
- What Physics experts to review work, but they are not familiar with our notation.
- Want to deliberately anchor experts towards high correlations between latent variables.
- Solution: Translate into natural language for expert review.
- Also, show table along with parameters so expert can see effect.

One parent case

```
fam <- PP.High$ForceAndMotion
PnodeRules(fam) <- "Compensatory"
PnodeLink(fam) <- "normalLink"
PnodeLinkScale(fam) <- sqrt(.2)
PnodeAlphas(fam)<-c(Physics=sqrt(.8))
PnodeBetas(fam) <- -.5
```

- Node Force and Motion: its only parent is physics.
- Link scale parameter only gives us R-squared, which is the percent of the explained variance by the predictors on the outcome variable. The value of R-squared is 0.8.
- There is a shift of about half of the standard deviation (.5) up towards more people having the skill.

Table 3: CPT for force and motion with physics

	Force and Motion		
Physics	High	Medium	Low
High	0.98	0.02	0.00
Medium	0.56	0.42	0.02
Low	0.04	0.52	0.44

Two parent case

```
eng <- PP.High$Energy
PnodeRules(eng) <-"Compensatory"
PnodeLink(eng) <-"normalLink"
PnodeLinkScale(eng) <- sqrt(.2)
PnodeLnAlphas(eng)<-
log(c(Physics=sqrt(.7),ForceAndMotion=sqrt(.9)))
PnodeBetas(eng) <- 0
```

- Node *Energy Can Transfer*: its parents are *Physics Understanding* and *Force/Motion*
- This skill has a compensatory rule, which means more of one parent offsets less of another.
- We are setting a regression of force/motion and energy on physics understanding. This model has two predictors naturally. Energy has both direct and indirect effect on physics understanding. Indirect effect goes through force/motion.
- One skill can offset another skill in this model. Link scale parameter only gives us R-squared, which is the explained variance by the predictors. The average value of R-squared is 0.8.
- The 0.8 comes from an average of physics=0.7 and force/motion=0.9.
- There is no difficulty shift. Somebody who is medium on the average of the parent variables will be roughly medi on the child variable.

The conditional Probability Table

Table 4: CPT parents of force and motion and physics

Physics	Force and Motion	Energy-High	Energy-Medium	Energy-Low
High	High	0.96	0.04	0.00
Medium	High	0.69	0.30	0.01
Low	High	0.21	0.66	0.13
High	Medium	0.62	0.36	0.01
Medium	Medium	0.17	0.66	0.17
Low	Medium	0.01	0.36	0.62
High	Low	0.13	0.66	0.21
Medium	Low	0.01	0.30	0.69
Low	Low	0.00	0.04	0.96

Experts reaction

- This expert didn't seem to think in this way.
- Expert mostly rubber stamped modeling team numbers.
  - Giving two alternatives might be better here.
- Intercept is probably easier to work with than difficulty (negative intercept).
- R^2 is a more natural parameter than the residual variance.
- Rescale slopes so that the average is 1.0 (let R^2 determine strength of relationship.

Outside-in validation of Bayes net

- Select a *Market Basket* of typical levels.
- Build evidence models for the market basket levels and attach them to competency model (*motif*).
- Enter typical student competency profiles and predict performance on market basket.
- Enter typical score patterns on market basket items and infer score profiles.
- Check with experts to make sure these are reasonable.
- *Next Step*: Do this in conjunction with pilot study.
- Market basket items can be release even if some levels are held back for security.

Lessons Learned

- Getting good variable definitions is both important and hard.
- Q-matrix is a useful tool for planning work.
- Presenting Experts with two choices is much better for getting a good discussion.
- Natural language translation is not enough: experts still need training.
- More natural parameterizations are useful.

Thanks!

- *Physics Playground v. 1*: Bill & Melinda Gates Foundation, grant *Games as Learning/Assessment: Stealth Assessment* (#0PP1035331, Val Shute, PI)
- *Physics Playground v. 2*: National Science Foundation grant *DIP: Game-based Assessment and Support of STEM-related Competencies* (#1628937, Val Shute, PI).

- Game Demo: `http://www.empiricalgames.org/ppunity5/LE/1.2.8/demo.html`
- Level Editor: `http://www.empiricalgames.org/ppunity5/LE/1.3.1/demo.html`
- Peanut and RNetica `https://pluto.coe.fsu.edu/RNetica`