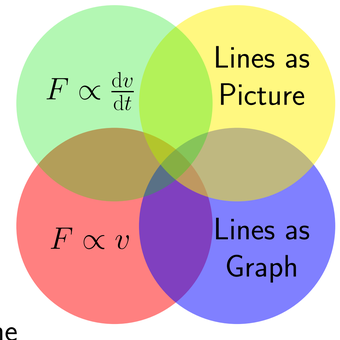
A multi-faceted approach to measuring student understanding

Ian T. Griffin, Nicholas J. Wright, Ryan Moyer, Kyle J. Louis, and Trevor I. Smith (smithtr@rowan.edu)

Department of Physics & Astronomy, and Department of STEAM Education, Rowan University, Glassboro, NJ 08028

Background & Assumptions

- Students use many different mental models to answer questions on the FMCE [1]
- Different questions and clusters are more or less conducive to particular models [3]
- Many students exist in a superposition state [2]
- Answers depend on both the student and the question



Normalized Gains & Model Analysis

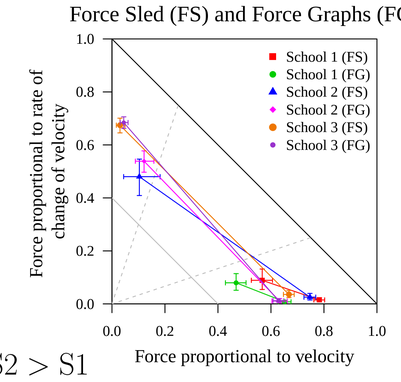
- **How much did my class learn?**
- **How much do their' ideas change after instruction?**

ANOVA results of individual student gains with Tukey HSD *post hoc* comparisons; * indicates $p < 0.001$.

	Average g			p-values			
	S1	S2	S3	Main Eff.	1v2	1v3	2v3
Full FMCE	0.29	0.60	0.69	*	*	*	0.003
Cases 1–4	0.23	0.61	0.71	*	*	*	0.02
Case 1	0.31	0.65	0.77	*	*	*	0.007

- All schools significantly different: $\mathrm{S3} > \mathrm{S2} > \mathrm{S1}$

Model analysis for three schools: Force Sled (FS) and Force Graphs (FG)



Models of Thinking

Model Name	Model Description	Cases
Correct, $F \propto dv/dt$	Consistent with Newton's 2nd law: net force is proportional to the rate of change of velocity.	1–4
Common, $F \propto v$	Net force is proportional to velocity.	1–4
Graph as Picture	Graphs can be interpreted as literal pictures of the situation.	1–4
$F \propto \lvert dv/dt \rvert$	Similar to the correct model, but ignoring sign/direction.	3,4
$F \propto \lvert v \rvert$; Graph Read Left	Net force is proportional to speed; reading the graph in the direction of motion.	4

Defining Cases

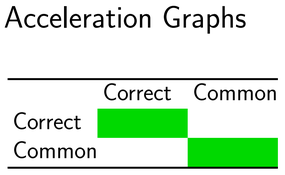
- **Do individual students answer coherently?**
- **How do their responses change?**

Identifying isomorphic questions

Case	Described Motion	Question FS	FG	AG
1	moving right, speeding up	1	16	22
2	moving right, steady speed	2	14	26
3	moving right, slowing down	3	18	23
4	moving left, speeding up	4	19	25

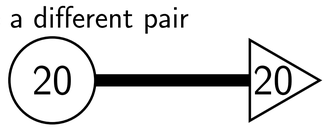
Contingency Tables

- Compare Force Graphs to Force Sled or Force Graphs to Acceleration Graphs
- Number of students who gave each response pair
- Diagonal cells show within-student coherent responses
- Large numbers show between-students consistent responses
- Cohen's w indicates the strength of the correlation between individual students' responses [4, 5]. weak: $w < 0.1$; moderate: $w \approx 0.3$; strong: $w > 0.5$
- Ignore models with fewer than 5% of responses on pre- and post-test



Consistency Plots

- Visualizing student transitions between table cells [6]
- "Arrows" show the number of students who went from one pair of pretest responses to a different pair
  - Start in circles (pretest)
  - End in triangles (post-test)



- Squares show students who did not change their answers
- Data sets may be compared using χ^2 test of independence; results are significant at $p < 0.05$

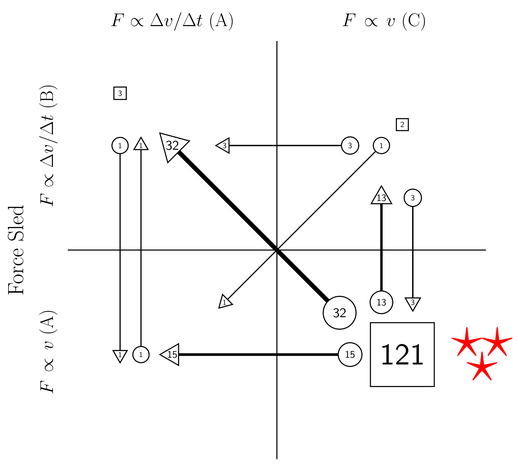
35

Case 1: Moving right, Speeding up; Force Sled vs. Force Graphs, All Schools

School 1

Force Sled \ Force Graphs	pretest $F \propto \frac{\Delta v}{\Delta t}$	pretest $F \propto v$	post-test $F \propto \frac{\Delta v}{\Delta t}$	post-test $F \propto v$
$F \propto \frac{\Delta v}{\Delta t}$	4	9	39	15
$F \propto v$	1	181	17	124
	$w = 0.48$		$w = 0.60$ ★★★★	

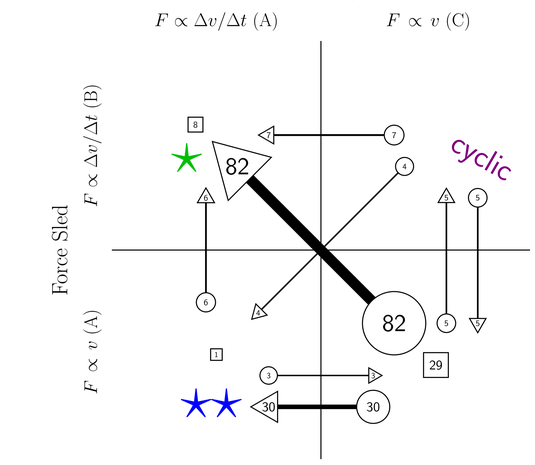
$N = 195$



School 2

Force Sled \ Force Graphs	pretest $F \propto \frac{\Delta v}{\Delta t}$	pretest $F \propto v$	post-test $F \propto \frac{\Delta v}{\Delta t}$	post-test $F \propto v$
$F \propto \frac{\Delta v}{\Delta t}$	8	16	103	5
$F \propto v$	10	146	35	37
	$w = 0.31$		$w = 0.54$ ★★★★	

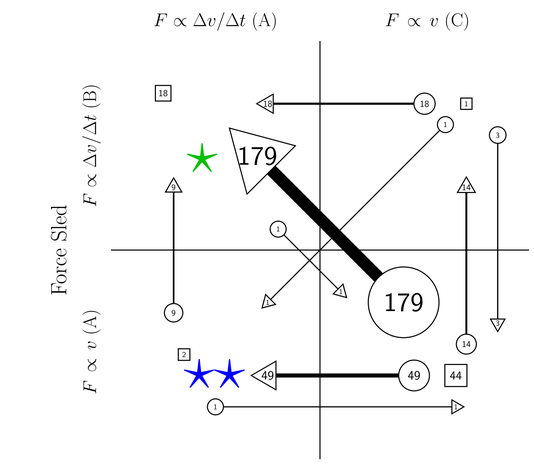
$N = 180$



School 3

Force Sled \ Force Graphs	pretest $F \propto \frac{\Delta v}{\Delta t}$	pretest $F \propto v$	post-test $F \propto \frac{\Delta v}{\Delta t}$	post-test $F \propto v$
$F \propto \frac{\Delta v}{\Delta t}$	19	23	224	15
$F \propto v$	12	286	52	49
	$w = 0.47$		$w = 0.49$ ★★★★	

$N = 340$



- Most students at Schools 2 and 3 go from common incorrect to correct on both (⋆)
- More students increase on Force Graphs than Force Sled (⋆⋆)
- Most students at School 1 stay in the common incorrect cell on all questions (⋆⋆⋆)
- Coherence increases over time (⋆⋆⋆⋆)
- Schools 2 & 3 visually and statistically similar; both are much different from School 1

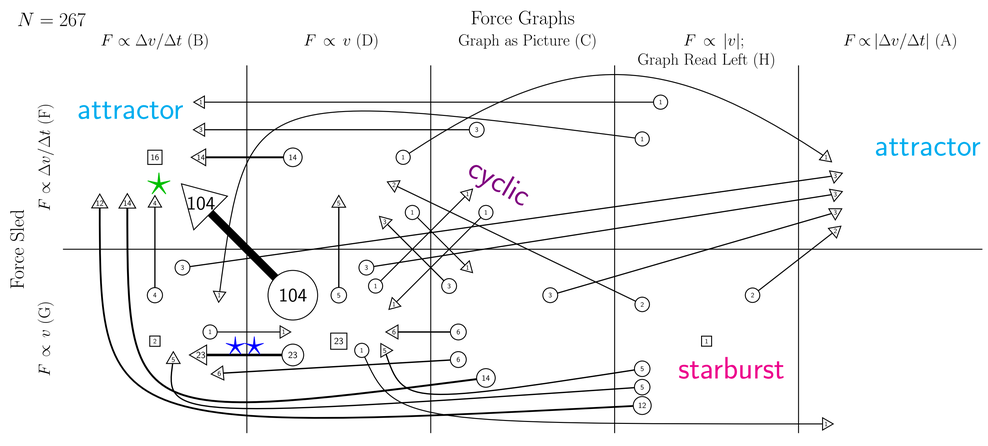
p-values from χ^2 test of independence: main effect significant at $p < 0.05$, pairwise significant at $p < 0.013$ (Bonferroni correction); * indicates $p < 0.001$.

	Main Effect	1v2	1v3	2v3
Case 1, FS/FG	*	*	*	0.24
Case 1, AG/FG	*	*	*	0.64

Case 4: Moving left, Speeding up; Force Sled vs. Force Graphs, School 3

School 3

Force Sled \ Force Graphs	$F \propto \frac{\Delta v}{\Delta t}$	$F \propto v$	Picture	$F \propto \lvert v \rvert$; left	$F \propto \lvert \frac{\Delta v}{\Delta t} \rvert$
pretest					
$F \propto \frac{\Delta v}{\Delta t}$	16	16	4	2	0
$F \propto v$	10	160	32	27	0
	$w = 0.45$				
post-test					
$F \propto \frac{\Delta v}{\Delta t}$	168	10	1	0	12
$F \propto v$	37	36	1	1	1
	$w = 0.52$ ★★★★				



- Many different transitions: "beginning state" + "instruction" ≠ "ending state"
- Possible hierarchy of incorrect responses [7]
  - Starbursts may represent very naïve responses (only pretest)
  - Attractors may represent more sophisticated ones (only post-test)
- Cyclic transitions only visible on consistency plots
- Similar statistical results for all cases

p-values from χ^2 test of independence; * indicates $p < 0.001$.

	Main Effect	1v2	1v3	2v3
Case 4, FS/FG	*	*	*	0.07
Case 4, AG/FG	*	0.004	*	0.33

Summary

- Different approaches reveal discrepant similarities and differences
- Model analysis explicitly treats students as being in a superposition state
- Contingency tables with Cohen's w show within-student coherence
- Consistency plots with χ^2 show transitions and reveal hidden similarities and differences

Future Directions

- Synthesizing results across cases
- Conducting interviews to test model definitions
- Developing statistic to report between-students consistency
- Closely examining similarities and differences between the instruction at each school

References

[1] R. K. Thornton and D. R. Sokoloff, Am. J. Phys. **66**, 338 (1998).
[2] L. Bao and E. F. Redish, Phys. Rev. ST Phys. Educ. Res. **2**, 010103 (2006).
[3] T. I. Smith, M. C. Wittmann, and T. Carter, Phys. Rev. ST Phys. Educ. Res. **10**, 020102 (2014).
[4] J. Cohen, *Statistical power analysis for the behavioral sciences*, 2nd (Lawrence Erlbaum Associates, 1988).
[5] R. Rosenblatt and A. F. Heckler, Phys. Rev. ST Phys. Educ. Res. **7**, 020112 (2011).
[6] M. C. Wittmann and K. E. Black, Phys. Rev. ST Phys. Educ. Res. **10**, 010114 (2014).
[7] R. K. Thornton, AIP Conf. Proc. **399**, 241 (1997).