Losing it: The Influence of Losses on Individual's Normalized Gains

K. Miller[1,2], N.Lasry[1,3], O. Reshef[2], J. Dowd[1], I. Araujo[1,4] and E. Mazur[1]

1 Harvard University, USA   2 McGill University, Canada   3 John Abbott College, Canada
4 Federal University of Rio Grande do Sul, Brazil

Kelly Miller   Nathaniel Lasry   Jason Dowd   Ives Araujo   Eric Mazur

Research Goals

Researchers and practioners routinely us the normalized gain (g), also known as Hake's Gain[a], to evaluate the effectiveness of instruction. Recently, concerns were raised about this metric because it implicitly assumes there is no loss (no correct pre-test responses become incorrect after instruction). Dellwo[b] proposes two alternative metrics to g for measuring conceptual change: Gain (G) and loss (L).

We use five years of introductory physics students' responses to the Force Concept Inventory (FCI)[c] to study the difference between the normalized gain and the new metrics. We determine the extent to which losses occur so that we may determine whether g provides misleading data regarding the conceptual gain that is taking place. Because Peer Instruction was implemented in these classes, we ultimately intend to determine whether the reported increased effectiveness of the physics instruction is confirmed by both metrics.

Hake's *gain*, *G and L*

To illustrate the role of losses we define four possible transitions that can occur between a pre-test and a post-test.

Pre-test	Pos-test	Transition
Right (*R*)	Right (*R*)	*RR*
Wrong (*W*)	Wrong (*W*)	*WW*
Right (*R*)	Wrong (*W*)	*RW*
Wrong (*W*)	Right (*R*)	*WR*

Hake's gain (*g*)	Gain (*G*)	Loss (*L*)
$g = \frac{WR - RW}{WW + WR}$	$G = \frac{WR}{W_{pre}}$	$L = \frac{RW}{R_{pre}}$

G is normalized with respect to the potential gain, and *L* is normalized with respect to the potential loss. However, in *g*, both the gain and loss are normalized with respect to the potential gain. Hake's gain implicitly assumes that the loss is zero. However, if this assumption is incorrect and RW>0, then Hake's gain contains a term that has no meaning.

Method

The FCI was administered at the beginning and end of the term during the first five years that Peer Instruction was being developed at Harvard University. The total student population size was N = 858. For each year of data, we recorded the number of transitions (RR, RW, WW, WR) for each student. From this, *g*, *G* and *L* were calculated for each student and yearly averages were calculated.

a) Hake, R. R. (1998). Interactive-Engagement vs. Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses. **Am. J. Phys.** 66(1): 64-74.

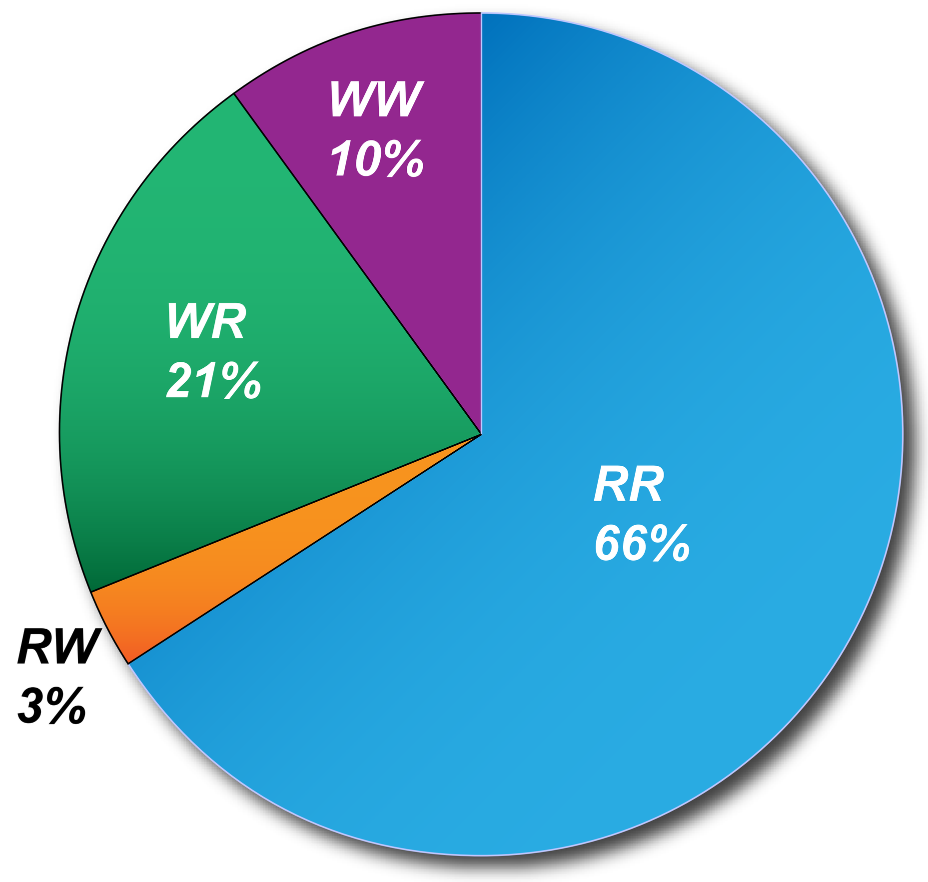
b) Dellwo, D. (2009). Reassessing Hake's Gain. Preprint, available on request.

c) Hestenes, H., Wells, M., Swackhamer, G. (1992). Force Concept Inventory. **The Physics Teacher** 30: 141-158.

Loss

The relative number of transitions for each of the four categories is shown. The percentage of answers changed from right to wrong, while small (3%), is not zero as Hake's gain assumes.

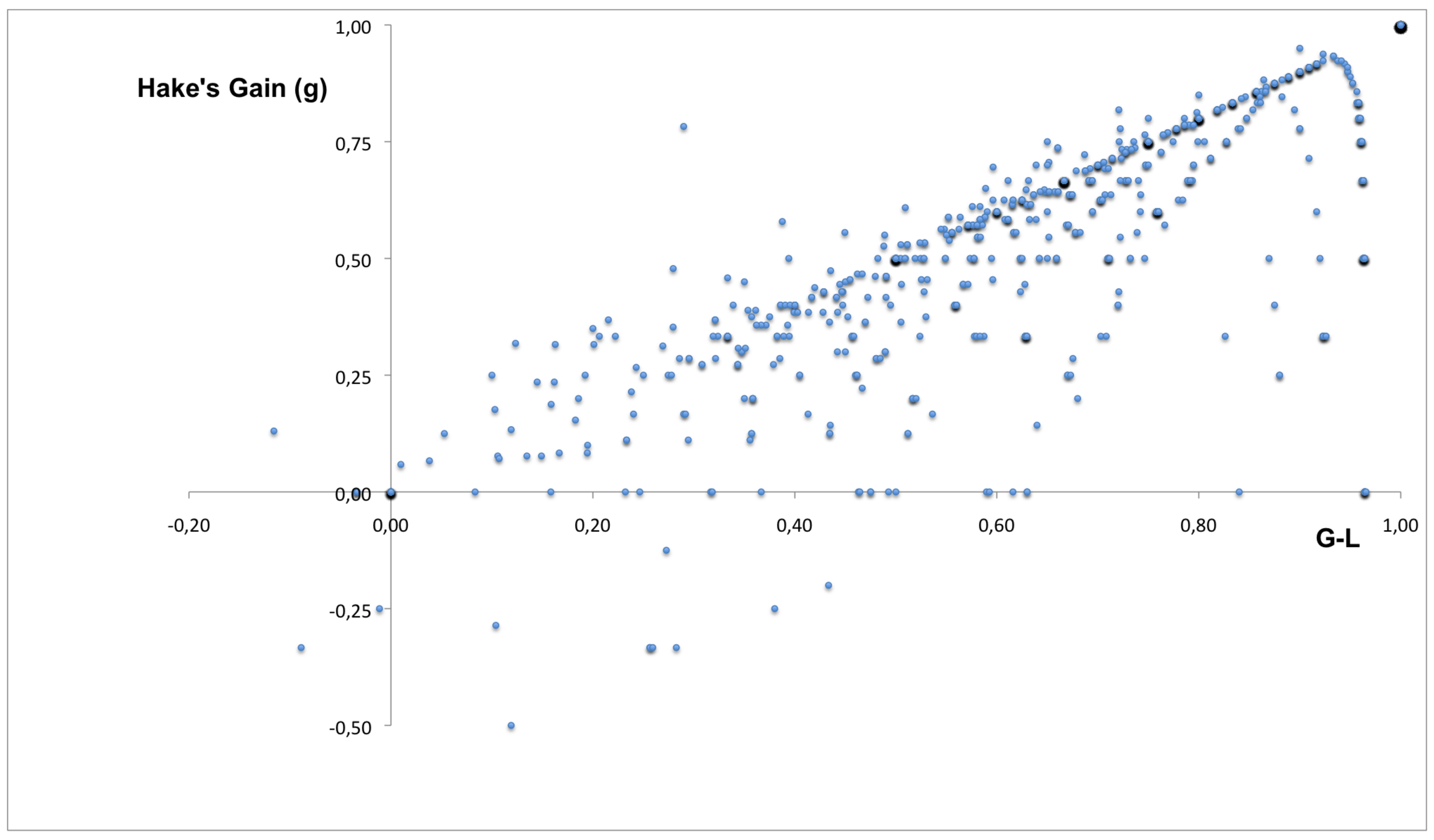
Quartile	*g*	Average *L*
1	0-0.4	12%
2	0.41-0.62	7%
3	0.63-0.8	4%
4	0.81-1.0	1%



FCI Average Transitions 1991-1996

Students with lower Hake's gains (first two quartiles) show a much higher percentage of losses than students with higher Hake's gains (last two quartiles).

g vs. *G-L*



This plot highlights the fact that Hake's gain is a good approximation of G-L at higher levels of g (when gain necessarily outweighs loss). A majority of the scatter occurs below the linear regression trend. For this particular student population, Hake's gain is generally smaller than G-L.

Conclusion

Using FCI data from Harvard University, we have shown that losses can be significant, especially for segments of the population whose Hake's gain is low. While determining both the G and L metrics is more computationally intense, they provide a more accurate depiction of conceptual change, particularly for populations whose Hake's gain is lower than 0.5. We will continue this analysis by comparing different student populations and pedagogies using G and L.

VE RI TAS

GROUP http://mazur-www.harvard.edu