

## Instructor Comments

Perhaps this question was framed too thinly. What I was trying to get you to think about was evaluation of a design as it is being developed and once released. Can we do this in a way that does not transform the experience but offers us confidence in the result? We might start, for instance, with the idea that Mitch presented last week: we want to create educational environments that are physical, manipulable and familiar. Now we are into the design process and we have a set of blocks that do something. How can we evaluate whether or not this design generates deep engagement? Iterative design, observation and focus groups are often used in combination to gain some insight into whether or not a design or particular content presented in a particular way works. Can we quantify how well a design works? Metrics might be: how many people just try it? are there bugs? how long do people continue to engage in one session? do people return or say they will return? can we understand the nature of the relationship people have with the thing we have made? In a course, an obvious metric is they learn the material well enough to use it in a new context? In the case of Introductory Physics at MIT, there were both perceived problems with the old design and a change in culture that may have influenced the new design. You have all developed attributes for engagement: some of you included concepts of duration in your set. Duration is hard to observe. What we would like to uncover is whether we can find some approaches to measurement that work across different problem spaces: a course, a robot? Dan and Roz will shed some light on what they think one can measure.