Research results for Mecanika, a game to learn Newtonian concepts

François Boucher-Genesse, Martin Riopel, Patrice Potvin, Université du Québec à Montréal,
francoisbg@gmail.com, riopel.martin@uqam.ca, potvin.patrice@uqam.ca

Abstract: A large body of research in mechanics indicates that interactive engagement teaching methods usually have higher chances of influencing students' conceptions than direct instruction. A few researchers specifically studied the impact of video games on Newtonian Physics instruction through empirical means, with some limited success. Mecanika is a free online game that sets itself apart from previous work by simply offering puzzling physics situations, without attempting to explain the theory in the game. Students who used the game as homework, facilitated with classroom debriefings and guidebooks, wielded significantly higher gain than a control group on the standard Force Concept Inventory test. Students who only played as homework registered a similar gain, even though Mecanika was never mentioned in the classroom. This gain was unexpected, since the game does not make any physics concept explicit, and was designed to be integrated in a classroom setting.

Mecanika trailer: www.youtube.com/watch?v=0yCTHV9Qv44
The game: www.gameforscience.ca/physica

The state of physics education

Many educators are advocating a qualitative and conceptual approach to understand Newtonians physics, which does not start with mathematical formulas, but rather with experiences, laboratories and demonstrations focused on students' conceptions (diSessa, 2001). Basing themselves on this large body of research, the Organisation for Economic Co-operation and Development recommends we make teaching physics more attractive, and focus on conceptions (OECD, 2008). These conceptions are often referred to as common sense intuitions based on observations made in everyday life. An example of a classic erroneous conception (hereby referenced as misconception) is to think that two balls of different weights, dropped at the same time from the same height, will hit the floor at different times.

Hestenes published a test that could reliably be used to assess whether students held conceptions that were Newtonians or erroneous: the Force Concept Inventory (Hestenes, Wells, & Swackhamer, 1992). The test is even recognized by its detractors to be the best available tool to assess mechanics teaching efficiency (Heller & Huffman, 1995). Perhaps the most interesting finding that followed is that traditional instruction (i.e., passive-student lectures, recipe labs, and algorithmic-problem exams) fail to convey much conceptual understanding of physics to the average student. Interactive-Engagement methods (i.e. methods designed at least in part to promote conceptual understanding through interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback through discussion with peers and/or instructors), however, were found to be much more successful (Hake, 1998).

Mecanika

A few researchers specifically studied the impact of video games on Newtonian Physics instruction through empirical means (Potvin & al, 2010; Rieber & Noah, 1997; White, 1984), although the most recent results were achieved by Clark & al (2010) with SURGE. We started to work on our own mechanics game, Mecanika, around the same time as work on SURGE began, and took a different approach by simply offering puzzling physics situations, without attempting to explain them explicitly in the game. The game also differentiates itself from the others by being a reflexive puzzle game: players do not have to react to quick events, and need to pause to predict the outcome of their actions.

Figure 1. A classic mistake in level B1 (backgrounds removed), which is related to the “last force to act determines motion” misconception

The game’s goal is to create a path of robots that will direct scouts over a set of stars. The scouts are produced by the top-left machine in Figure 1, and are basically inert boxes. In this simple introduction level, for example, you start the level in a zero-gravity environment with a punching robot already placed at the exit of the machine, which will give scouts an impulse in the right direction. Players have to place another impulse robot in the level, which will give an equally powerful hit downward. Most students will at this point place the impulse robot directly over the second star, expecting the scout to move in the Y axis only, as shown in Figure 1. This is a misconception that Hestenes identified as “CI3 - last force to act determines motion”. Players will eventually realize that both impulses have an impact on the scout’s direction, and place the punching robot over the first star to reach the second one (see Figure 2).

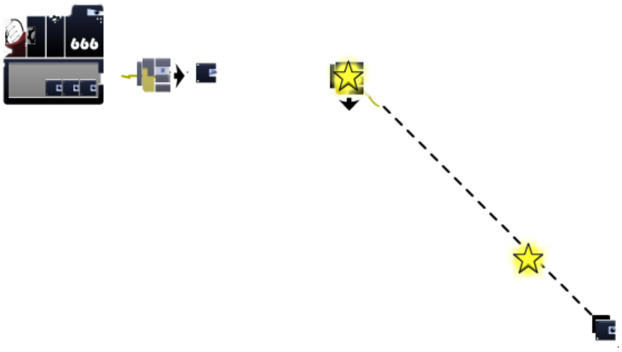

Figure 2. The solution to level B1

Mecanika features 50 levels, each focusing on misconceptions identified by Hestenes. Players place robots that generate impulses, continuous force areas, circular movement or even toggle gravity. Since the game is designed to be played as homework, considerable effort was spent on the production value of the game to make it a compelling activity for students at home. The game was developed by researchers at the University of Quebec in Montreal and the Creo Montreal game studio over the last two years. It contains 3 to 4 hours of gameplay, and is available for free at www.gameforscience.ca/physica.

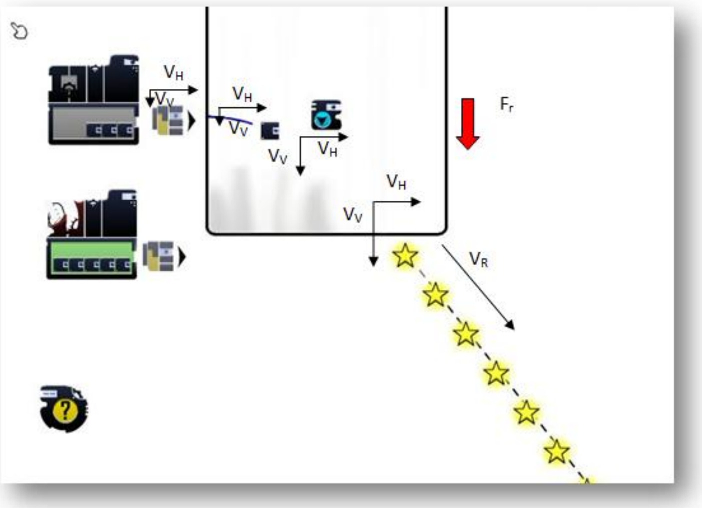

Figure 3. A sample illustration from the teacher's guidebook

Formal understanding of the mechanics concepts happens in the classroom, where teachers use detailed pedagogical guidebooks (100+ pages, see Figure 3) to explain what students intuitively learned in the game. Students also have to describe why they think their own puzzle solutions worked. These guidebooks are available to teachers and researchers, along with videos that explain the material behind each level, on a teacher's portal. Access to this portal is restricted, but will be granted if you email francoisbg@gmail.com from a school/university email. An English version of the guidebooks will be available by the end of summer 2011.

Research methodology

The game was studied in real classroom environments, in order to benefit from any instructional support that could occur there (O'Neil, Wainess, & Baker, 2005). The research methodology can be seen in Figure 4. Two teachers, each with four classrooms, participated in the study. The first part of the experiment is a typical experimental/control group setup with post and pretests. Each teacher first had their students take the Force Concept Inventory test, and taught as they would regularly for two of their classes. The two other classes got the same instruction, from the same teacher, but also played Mecanika as homework. They then filled out their student's guidebooks, and teachers debriefed them in the classroom about their game experience. Finally students from all groups took the FCI test again as a posttest.

One teacher used the game over one month; the other used the game sporadically throughout the term over a three months period. The overall time spent on the game, guidebooks, and classroom debriefings is about the same for both teachers, and they both used about the first two thirds of the game's levels.

Figure 4. The research procedure used to study Mecanika. Two different ways to use the game in the classroom were studied, each time comparing to a control group.

In the second part of the experiment, students that played the game stopped using it, and continued with regular instruction. The players that did not yet play the game played it as homework, but did not receive paper guidebooks to fill, and did not benefit from classroom debriefing. They were only told to play the game as homework over a one month period. Every student then took the FCI test again.

Results and discussion

Impact with classroom debriefings

In the first part of the experiment, the control group did not see a significant increase in their overall FCI score (p=0.08, +1.9%, effect size d=0.19, N=82), but the experimental group had a significantly different gain (p<0.001, +9.2%, effect size d=0.95, N=51). The changing variables between the experimental and the control groups are the inclusion of Mecanika and guidebooks as homework, and the game discussions that happened in the classroom. This is an important result, since most game studies using a control group end up with similar results between the two groups (Hays, 2005).

The effect size is measured using Cohen's d, and can be considered to be a "large" effect (over the 0.8 threshold). But a perhaps more telling way to assess if the game caused a significant gain would be to look at other instruction methods that were studied using the same FCI test. One such study was conducted by the authors of the Force Concept Inventory in a nation-wide experiment called the Modeling Instruction Project. The researchers designed "an intensive 3-week Modeling Workshop that immerses [teachers] in modeling pedagogy and acquaints them with curriculum materials designed expressly to support it" (Hestenes, 2006). 66 teachers participated in this experiment (N=3394), which was conducted over a full term. As illustrated by Figure 5, the teachers which participated in the modeling workshop registered an important gain over the term, a gain which was 10% higher than the control groups.

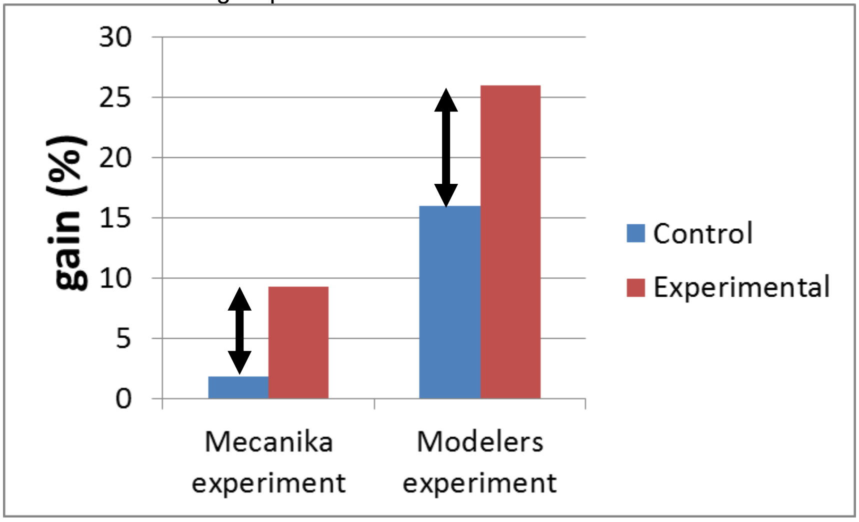

Figure 5. Gains for the Mecanika and Modelers Instruction Project experiments. Both gains, when compared to their control groups, are of comparable size.

The Mecanika experiment was in comparison much shorter, but still produced a difference in gain between the experimental and control group of 7.4%. What is further interesting is that very limited training was given to the teachers: no more than 30 minutes was spent talking about the game in person. The results by no mean indicate that we should give the game to teachers instead of training them properly, but they do point to the possibility of rapidly enhancing students' Newtonian conceptions just by giving the game and pedagogical guidebooks to teachers across the country.

We were able to gather how much each student has played through Mecanika, and could thus observe which portion of the game seemed to cause a more important FCI gain (see Figure 6). The first ten levels were used to teach game mechanics, which would explain why no significant increase was found between groups. Levels 20 to 30 also did not seem to have much of an influence on FCI items. A potential explanation could be that although these levels contained situations similar to the ones seen in the FCI, the game and the test contexts were different.

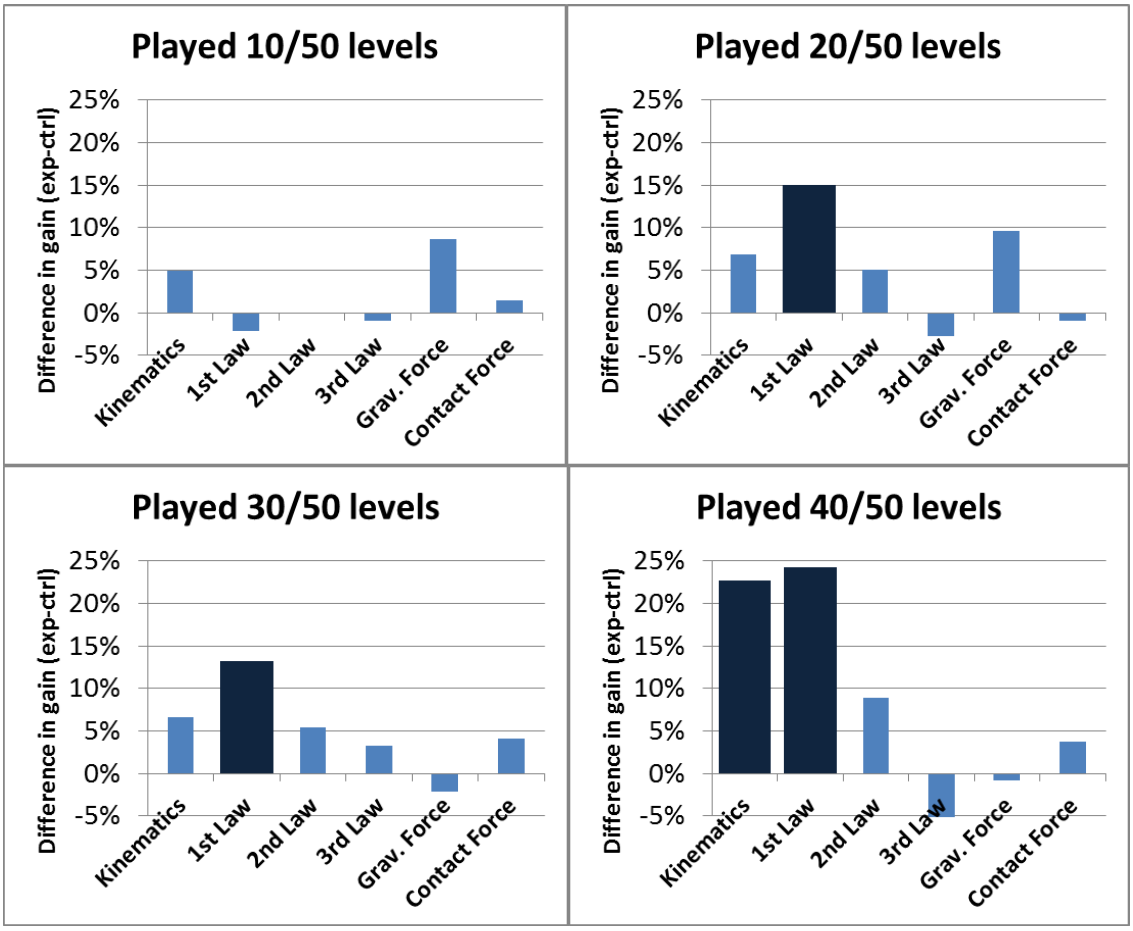

Figure 6. Difference between the experimental and the control group gains, categorized by concept. Students have been separated by their progression in the game. Larger and darker columns represent a difference in gain which was significantly different (p<0.05).

By dividing the FCI items in categories, the game's impact can also be studied more closely (Savinainen & Scott, 2002). Most of the overall test gain can be attributed to increases on Newton's First Law and on Kinematics FCI items. The game should thus be used when talking about these principles in the classroom. The focused impact was to be expected, since the game design didn't target all concepts covered by the FCI; the game's scope had to be limited in order to keep a consistent game design throughout all levels. The remaining levels at the end of the game were designed to focus on other concepts, but were not tested.

Impact without classroom debriefings

By looking at the experimental setup, one could wonder if the gain really happened because students played the game, or because they had guidebooks and classroom debriefings about it. The second part of the experiment can shed some light on this matter, since students that didn't play the game yet played it later as homework only. Teachers had explicit instructions not to talk about Mecanika in their classrooms. Since the game never explains concepts clearly, or even names the observed situations, the hypothesis was that the game by itself would have a much smaller effect than if the guidebooks and classroom debriefings were used as well. Not having debriefings also meant that the students would play for about 1.5 hours at some point in the month. We further thought that measuring an increase in gain for such a short activity over a full month would be much harder. We were surprised to see that the students got a gain (p=0.02, +7.3%, effect size d=0.59, N=26) which is not significantly different from the students which had guidebooks and classroom debriefings. The lower number of participant is explained in the limitations section.

One could look at these results and make the hypothesis that the gain is due to the teachers changing their instructions methods. If that was the case though, we would argue that we would also see an increase in the control groups, which was not the case. We are left to guess that either or both of these following hypotheses can explain the relatively high gains: 1 – the guidebooks were poorly designed, or the classroom debriefings could have been done better, or 2 – most of Mecanika’s potential comes from just playing with it.

Additional findings

Girls and boys did not get a significant gain difference, but when asked, boys did think that the game and the guidebooks were more useful, and that the game was more fun ($p<0.05$, effect size ranges from $d=0.43$ to $d=0.49$). It is also interesting to note that the gain registered by the experimental group in the first part of the study was left virtually intact one month after. The concepts were retained and no significant difference was observed during the last month of traditional instruction ($p=1.00$, +0.0%, effect size $d=0.00$, $N=55$).

Limitations

The previous statistics had a low amount of students in experimental groups. The reasons are twofold. First, some students were not able to play since the game was at times lagging too much. Mecanika is integrated in a larger flash MMO-like world, www.gameforscience.ca, which at the time slowed down considerably when more than twenty people joined in simultaneously. This bug, combined with the fact that Mecanika is a pretty heavy flash game, made it not playable for many: 48% of students said they had technical problems that prevented them from playing at some point. The second reason that could explain a lower-than-expected participation rate is that play was made mandatory by teachers, but wasn’t reinforced by making the results count on their class score for example.

Another important limitation to this study lies in that only two teachers participated to the study, despite the relatively large number of students. More teachers would have allowed us to see if other ways of debriefing on the game in the classrooms could have resulted in higher gains. The two teachers we had were also not randomly selected – they were recruited for their interest in the project. We should add though that one of them was not acquainted with technology, and obtained similar gains to the second teacher, which played games regularly.

Conclusion

Multiple interesting research avenues remain, such as investigating if we could train teachers to make a better use of the game, or doing A/B testing to investigate the impact of some game mechanics on learning. These research questions can be easily answered, since we now know that the game will most likely have a measurable impact. The research team is open to share the game and resources with other teams in order to investigate these questions, and can be reached by using the contact information on this paper.

Mecanika will be publicly launched in the Fall of 2011, and is mostly finished at this point. Much design insight was gained from studying the learning results from the game, and the company behind the game’s graphics and story, Creo, is now looking for funding on the second iteration of the game.

The findings presented in this paper make it clear that even a low involvement on the part of teachers, by giving the game to play as homework, helps transform the students’ conceptions into Newtonians conceptions. Whether or not other means of using the game in classrooms, computer laboratories or at home could wield higher results is still an open question.

References

Clark, D., Nelson, B., D'Angelo, C., Slack, K., & Garza, M. (2010). SURGE: integrating Vygotsky's spontaneous and instructed concepts in a digital game? In *Proceedings of the 9th International Conference of the Learning Sciences - Volume 2* (pp. 384-385). Chicago, Illinois: International Society of the Learning Sciences. Retrieved from http://portal.acm.org/citation.cfm?id=1854509.1854709

diSessa, A. (2001). *Changing Minds: Computers, Learning, and Literacy*. Cambridge: The MIT Press.

Hake, R. R. (1998). Interactive-engagement vs Traditional Methods in Mechanics Instruction. *American Journal of Physics*, *66*(1), 64-74.

Hays, R. T. (2005). *The Effectiveness of Instructional Games: A Literature Review and Discussion*. Retrieved from http://stinet.dtic.mil/oai/oai?&verb=getRecord&metadataPrefix=html&identifier=ADA441935

Heller, P., & Huffman, D. (1995). Interpreting the force concept inventory: A reply to Hestenes and Halloun. *The Physics Teacher*, *33*(8), 503.

Hestenes, D. (2006). Notes for a Modeling Theory of Science, Cognition and Instruction. Presented at the GIREP conference: Modelling in Physics and Physics Education.

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

OECD. (2008). *Encouraging student interest in science and technology studies*. OECD Publishing.

O'Neil, H. F., Wainess, R., & Baker, E. L. (2005). Classification of Learning Outcomes: Evidence from the Computer Games Literature. *Curriculum Journal*, *16*(4), 455-474.

Potvin, P., Riopel, M., Charland, P., Ayotte, A., & Boucher-Genesse, F. (2010, October). *Enhancing performance in the "Force Concept Inventory" test using homework gameplay while involving physics teachers in the level design process : SpaceFart*. Presented at the Meaningfulplay 2010, East Lansing, MI.

Rieber, L., & Noah, D. (1997). Effect of Gaming and Visual Metaphors on Reflective Cognition Within Computer-Based Simulations. Presented at the AERA conference, Chicago, IL. Retrieved from http://it.coe.uga.edu/~lrieber/gaming-simulation/

Savinainen, A., & Scott, P. (2002). Using the Force Concept Inventory to monitor student learning and to plan teaching. *Physics Education*, *37*(1), 53-58.

White, B. Y. (1984). Designing Computer Games to Help Physics Students Understand Newton's Laws of Motion. *Cognition and Instruction*, *1*(1), 69-108.

Acknowledgments

The researchers would like to extend their special thanks to Creo, without which Mecanika would have never reached the esthetical level, production value and users it currently has. They would also like to extend their thanks to the Inukshuk organization, which funded the first iteration of the game, allowing it to be available online for free in both French and English.