Learning Qualitative Relations in Physics with Law Encoding Diagrams

Peter C-H. Cheng
ESRC Centre for Research in Development, Instruction and Training.
Department of Psychology, University of Nottingham,
Nottingham, NG7 2RD, U.K.
peter.cheng@nottingham.ac.uk

Abstract

This paper describes a large scale experiment that evaluates the effectiveness of Law Encoding Diagrams (LEDs) for learning qualitative relations in the domain of elastic collisions in physics. A LED is a representation that captures the laws or important relations of a domain in the internal structure of a diagram by means of diagrammatic constraints. The subjects were 88 undergraduate physics students, divided into three learning trial conditions. One group used computer based LEDs, another used conventional computer based representations (tables and formulas), and the third was a non-intervention control group. Only the LED subjects had a significant improvement in their pre-test to post-test qualitative reasoning. The LEDs appear to make it easier for subjects to explore more of the space of different forms of collisions and hence gain a better qualitative understanding of the domain.

Introduction

Law Encoding Diagrams (LEDs) are an interesting class of diagrammatic representations. They capture the law(s) of a domain by means of their internal structure, using geometric, topological or spatial constraints, such that each diagram represents a single instance of the phenomenon or one case of the law. LEDs can be found in the history of science (Cheng, in press) and may have had an significant role in some discoveries, such as finding the law of conservation of momentum (Cheng & Simon, 1992). As LEDs seem to have been useful to original scientists, it is possible that they may help students learn about the same domain. Cheng (1994, 1995) describes a detailed small scale study in which subjects learnt about elastic collisions in physics. The subjects used computer based LEDs in a system called ReMIS-CL. It was found that subjects, physics students, could quickly learn to use LEDs for problem solving with little instruction. In post-tests half the subject used LEDs for problem solving with novel strategies, in contrast to their own ineffective pre-tests solutions. From the detailed analysis of their use of ReMIS-CL, it appears that the successful subjects obtained a better understanding of the diagrammatic constraints of the LEDs, because they comprehensively examined the space of structural forms of the LEDs.

This paper describes a larger scale investigation of computer based LEDs for learning, with 88 undergraduate physics students. The main aim was to evaluate the effectiveness of ReMIS-CL against controls of two kinds: (i) a group using a similar, but non-diagrammatic, computer based learning environment; and (ii) a non-intervention group. The investigation also provided further evidence to support the hypothesis that successful learning with LED is linked to the extent to which subjects explore the space of different structural forms of LEDs.

The next section of the paper describes the domain of elastic collisions and ReMIS-CL. The method and results are then outlined in the following two sections. The implications of the results are described in the final two sections, which also contrasts the present approach to others in computer based physics learning.

Elastic Collisions and ReMIS-CL

Elastic collisions are important in physics, because both momentum and energy conservation are involved. Here impacts between two bodies (balls) travelling in a straight line are considered. Figures 1 and 2 shows screen displays (minus menu bar) of ReMIS-CL, a computer based discovery learning environment for this domain. At the bottom of the screen, there is an animated simulation of the collision that the user can run at will. The two large areas above have two interactive LEDs: the *one-dimensional property diagram* (1DP diagram) and the *velocity-velocity graph* (VV graph). Figures 1 and 2 show different collisions but LEDs within each represent the same collision. The lines in the diagrams represent magnitudes of velocities and masses: $U1$ and $U2$ are the velocities before impact; and, $V1$ and $V2$ are the velocities after collision. In Figure 1 the bodies approach and depart in different directions but with equal speeds. In the 1DP diagram, mass lines, $m1$ and $m2$, are drawn equidistant between the $U1-U2$ and $V1-V2$ lines. In the V-V graph the masses are represented by the sides of the small triangle. The ratio of the lengths of the mass lines in both LEDs equals the ratio of the masses of the two balls.

Both LEDs can be directly manipulated to change the values of the variables. Figure 2 shows the result of sliding the handle, the small rectangle, at the end of the $U2$ arrow to the right, doubling the initial speed of body-2. The rest of the 1DP diagram's structure is automatically updated to be consistent with its own diagrammatic constraints and thus to satisfy both conservation laws. The LEDs are inter-linked so that the structure of V-V graph is also revised. ReMIS-

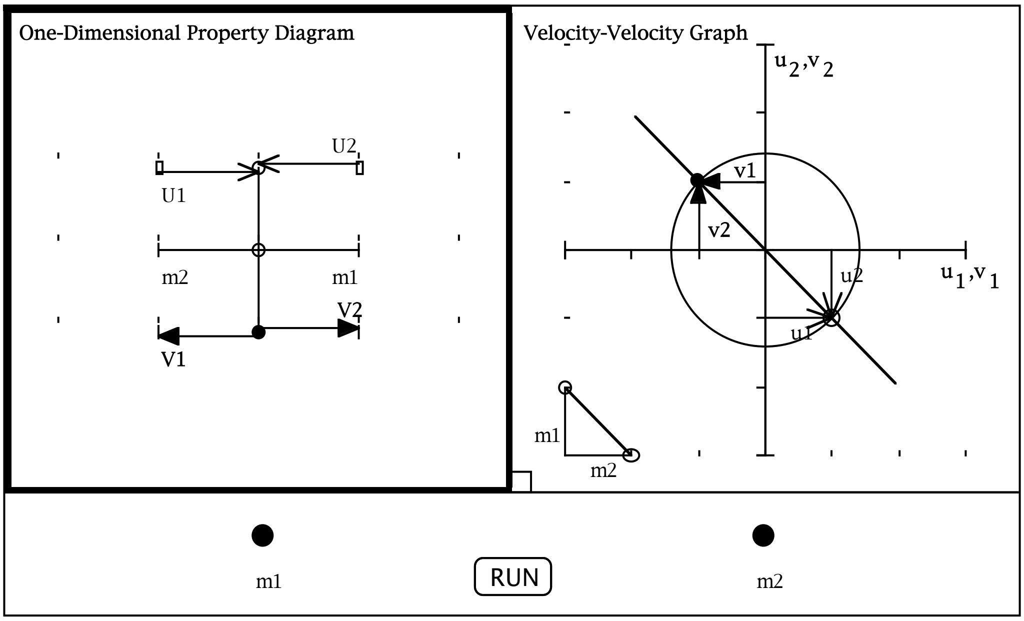

Figure 1: ReMIS-CL Screen showing simple equal mass collision.

CL ensures that the LEDs are always consistent with their diagrammatic constraints.

The three main constraints of the 1DP diagram are: (i) the tail ends of the arrows for the initial velocities and the points of the corresponding final velocity arrows must be in line vertically, making the total length of the $U1-U2$ line equal to that of the $V1-V2$ line; (ii) the total length of the mass line equals the length of the velocity lines; and, (iii) the ends of the lines not previously fixed in (i) and (ii), indicated by the small circles, must lie on a straight vertical or diagonal line.

In the V-V graph the straight diagonal line is constant momentum contour and the circle (sometimes an ellipse) is a constant energy contour. There are 3 main constraints. (i) The momentum contour line passes through the points for initial and final velocities, as indicated by the small circles, and is parallel to hypotenuse of the mass triangle. (ii) The centre of the energy circle/ellipse is at the origin of the

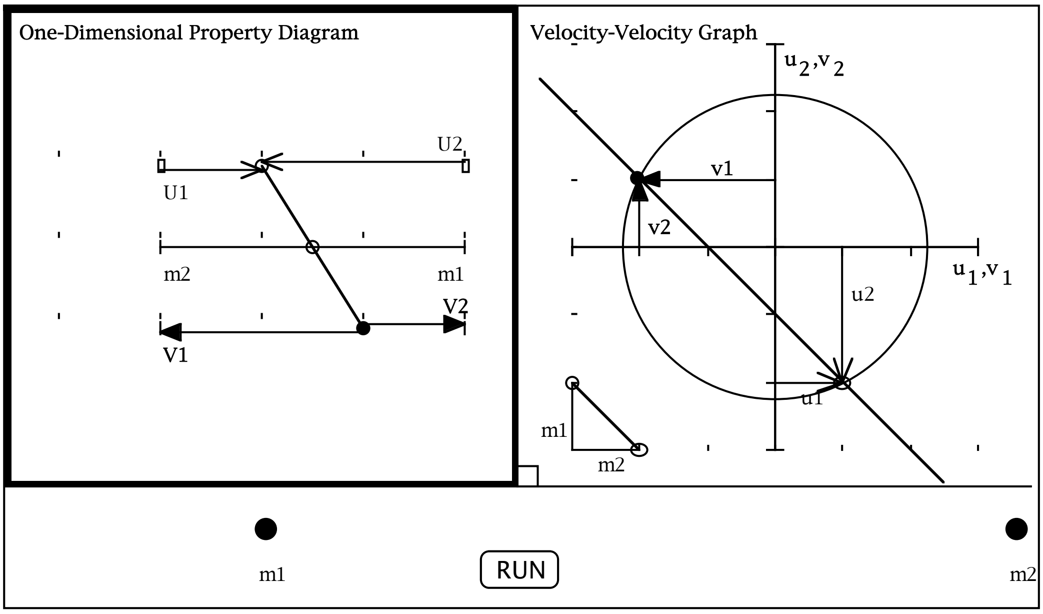

Figure 2: An Asymmetric collision in ReMIS-CL

graph and it also passes through the points for initial and final velocities. The intersections of the contours give the solutions to the two conservation laws. (iii) The eccentricity of the ellipse is given by the square-root of the ratio of the masses; $\sqrt{(m1/m2)}$. The law encoding constraints of this LED are more complex than those for the 1DP diagram.

ReMIS-CL logs the actions of the users on the system and stores the values of variables. Different kinds of operations are recorded with time stamps, including: switches between the representations; changes to values of the variables; resets that return the variables to their default values (i.e., Figure 1); and runs of the simulation.

One way to distinguish different collisions is in terms of *configurations*, defined as relations between the pairs of initial (or final) velocities with respect to their signs (directions) and whether they are zero, equal or unequal (Cheng, 1995). There are 7 different configurations for pairs of initial velocities, *U-configurations*. (The same applies for final velocities.) Figures 1 and 2 show different configurations. For different ratios of the masses, alternate final velocity configurations can result from each U-configuration. Combining initial and final velocity configurations, there are 41 possible *complete* configurations. Each corresponds to different structural forms of the LEDs, so this provides a convenient way to record and assess the behaviour of subjects.

Method

Design. There were three experimental learning trial conditions. (1) The LED group used ReMIS-CL with the two LEDs. (2) The Num group used numerical/formula version ReMIS-CL which had the animated simulation but not the LEDs. In their place were: (i) a temporally ordered table of previous sets of values from which particular cases could be simply selected; and, (ii) a structured table of current values of the variables along with values of momentum and energy terms for each body before and after collision. The values of the variables could be changed by selecting a variable and typing a new value. (3) The Con group was a non-intervention control group, so did not use either version of the system.

Subjects. The subjects were first year undergraduate physics students at the University of Nottingham. They participated in the experiment during a weekly computer programming class. Given the constraints on the organization of the class, which was run in six separate groups, it was not possible to randomly assign subjects to the three experimental conditions. Pairs of class groups made up each of the three experimental groups.

Materials and Procedure. The experiment included: a pre-test of the subjects knowledge of the domain; a learning trail on the system (not with the Con group); and, a post-test similar to the pre-test. The trial immediately followed the pre-test and the post-test followed a fortnight later. Pre-test and post-tests lasted 20 minutes and the trial 30 minutes.

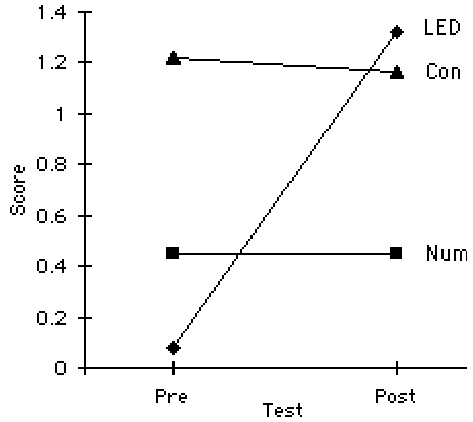

Figure 3: Mean Scores

The pre-test had four sections. The first asked the subjects to state the laws and write the equations governing elastic collisions. The second section had three qualitative questions in which the one or more outcomes of particular collisions had to be given. For example, what are the possible outcomes when the balls have different masses but they approach from opposite directions with the same speed? The third section included two questions about extreme cases, in which outcomes of collisions with large speeds or masses were considered. The final section was a quantitative question in which an exact solution had to be calculated. The questions were written in English and accompanied by a simple diagram depicting the situation (not a LED). Subjects responded by making annotated sketches. Most subjects had insufficient time to complete the last (quantitative) section, so it will not be considered here. The post-test was similar to the pre-test, except the first section was replaced by a general question about what they had thought they had learnt using the system. The LED and Num groups were also given a picture of the ReMIS-CL interface they had used.

During the trials the LED and Num groups worked in pairs. They were given a sheet that described how to use the system. For the LED group this included a brief description of the 1DP diagram and the V-V graph. The subjects were told to use the system to find out as much as they could about elastic collisions and given a log sheet on which to record their discoveries. ReMIS-CL logged the actions performed by the subjects.

For the purposes of analysis, only those subjects who did both the pre-test and post-test are considered. The groups numbers were: LED=25, Num=27 and Con=36.

Results

The subjects had reasonable conceptual understanding of the domain, with 97% and 76% knowing that momentum and energy conservation laws, respectively. There were no apparent differences among the three experimental groups in this respect.

Group	Operations	Proportion of simulations	Ratio of simulation series	Proportion of value changes	Ratio of value change series
LED	132	0.26	0.040	0.56	0.13
Num	52.5	0.44	0.092	0.47	0.093
LED/Num	2.51	0.60	0.44	1.21	1.37
p (*t* test)	<.001	<.001	<.001	<.001	<0.01

Table 1: Numbers and Types of Operations

Figure 3 contrasts the average pre-test and post-tests scores for all three groups on the qualitative problems. The scores were the sum, for all three questions, of the number of correct outcomes less the number of incorrect outcomes. The lowest and highest scores observed were -5 and +5. There was a significant increase in the LED group's score pre-test to post-test (t= 2.66, p=0.014), but the score for Num group was unchanged and the slight decrease of the Con group was not significant (t=0.80, p>.1). The difference in the pre-test scores of the LED group and the Con group was significant (t=2.4, p<.05). In a 3X2 mixed ANOVA of the 3 groups and the 2 tests, there were no significant effects of group or time of test, but there was a significant interaction ($F_{2,85}$=3.66, p<.05). There were strong and significant correlations between pre-test and post-test scores of the Num and Con subjects (Pearson; r=0.75, p<.001, and r=0.62, p<.001, respectively), but the LED group had a weak (non significant) correlation (r=0.17, p>.05). The LED group's qualitative reasoning has improved. The Num and Con groups' have not, and it appears that there has been little change in performance at the level of individual subjects.

Table 1 shows various average measures of the behaviour of the LED and Num groups in the trials. The LED group performed 2.5 times more operations than the Num group. Values changes and simulation runs were the most common operations. However, as a proportion of the number of operations of each subject, the LED group did significantly more value changes and ran fewer simulations than the Num group. Num subjects often performed series of two or more consecutive operations of the same kind; e.g., running the simulation three times in a row, with the same values. The ratio of the numbers of such series of operations to the total number of operations, per subject, are also shown in Table 1. For example, on average, 13% of operations by LED subjects were changes to the values followed by a series of at least one more change in the values. The LED group had significantly fewer simulation series and significantly more series of changes of values than the Num group. The LED group had three times the ratio of value change series to simulation series, whereas the Num group had nearly equal ratios. This implies that the two groups may be doing different kinds of reasoning, whilst using the system.

Simple measures of the distribution of U-configurations and complete configurations were devised, such that unity indicates an even distribution in which there are equal numbers of all the different configurations and zero indicates a maximally skewed distribution in which only one configuration is present. Larger values (approaching 1) indicate that a subject has explored more of the space of configurations. Table 2 shows the scores for the two groups, obtained by analysing the logs of subjects on ReMIS-CL. Both scores for the LED group were higher than those of the Num group, although only the difference in the complete configuration scores was significant. LED subjects are exploring a greater range of the possible configurations of collisions.

Group	U	Complete
LED	0.56	0.30
Num	0.47	0.20
LED/Num	1.19	1.5
p (*t* test)	.1>p>.05	<.001

Table 2: Measures of configuration distributions

Table 3 shows correlations for comparisons of U-configurations with other measures of performance on the systems. The LED group seems to have more thoroughly explored the space of different configurations by concentrating on value changes as opposed to runs of the simulation, shown by the significant positive correlations between the U-configuration distribution and value changes, both proportion and series measures. There are weak negative correlations for simulation measures and U-configurations

	Complete configuration distribution	Proportion of simulations	Ratio of simulation series	Proportion of value changes	Ratio of value changes series
LED	0.86	0.19	-0.06	0.62	0.68
p	<.001	>.05	>.05	<.01	<.001
Num	0.79	-0.33	-0.30	0.23	0.18
p	<.001	>.05	>.05	>.05	>.05

Table 3: Correlations between U-configuration distribution and other trial measures

for the Num group, which implies that running more simulations coincides with less exploration of the space of configurations.

The results for the two extreme problems were similar, with a significant improvement in LED subject's score pre-test to post-test, but no significant improvements in either the Con or the Num groups. However, the details of the results of the extreme case questions are consistent with the qualitative questions results, they add little to the interpretation of the results (in the next section), so are not reported here, due to limited space.

Discussion of Results

The experiment has shown that LEDs implemented as interactive diagrams can be effective for learning about qualitative relations and extreme cases in the domain of elastic collisions. This outcome is noteworthy, because the duration and form of the intervention. The subjects had a short time on the system, 30 minutes. They were not given a carefully designed series of activities, but merely told to look for interesting relations or patterns. As the LED group was the only one to improve, it appears that the gain can be attributed to the role of the LEDs. Alternative explanations will be considered and shown to be implausible.

Comparison of the Num and LED groups shows that the mere use of a computer based system with an animated simulation is not responsible for the gain. The Num group used the simulation more than the LED group. The difference between the pre-test scores of the Con and the LED groups was likely due to the non-random assignment of the subjects to the experimental conditions, which was beyond the control of the experimenter. Thus, it is possible to argue that the effect is due to repeated testing, if two assumptions are made. First, the Con group are experiencing a ceiling effect, so repeated testing will not result in an improvement. Second, the Num group does improve due to repeated testing, but this is matched by an equal degradation due to the use of the numerical/formula version of the system, for some unknown reason. This alternative is less likely than the explanation that the improvement is due to the LEDs, because it is more convoluted and the assumptions are not secure. First, it is unlikely that the Con group is experiencing a ceiling effect, because their average score of 1.2 is much less than the maximum qualitative problem score of 5. The second assumption is also weak, because it requires two independent processes simultaneously working to produced effects that cancel each other out. The strong and highly significant correlation between the Num subjects' pre-test and post-tests scores suggests that no processes are at work, rather than two independent ones.

The improvement in the LED group's score from significantly below to just above the Con group might be explained by contending that the LED group was more highly motivated, because they knew less. This is unlikely given the lack of change in the Num group and the small absolute (though significant) difference between the LED and Num pre-test scores.

Further, there are good reasons to positively attribute the improvement of the LED subjects to the LEDs. Part of the explanation of why the LED group improved, but the Num group did not, may be the sheer numbers of operations that the LED group did compared to the Num group, two and a half times as many. This difference can itself be explained in at least two ways. First, the LED version of the system is easier to use; a human computer interaction factor. The 1DP diagram and the V-V graph can be quickly changed by re-sizing a line. With the numerical/formula system it was necessary to select the variable and then type the required value at the keyboard, a slower process. However, this operational factor is not sufficient to explain the whole difference.

The second explanation, which is itself an explanation of the relative success of the LED group, is in terms of differences in the types of reasoning that the LED and Num subjects were doing. The Num subjects appear to have been spending time relating the values of the variables to the simulations, in the cases they examined. The Num group's greater proportion of simulations and ratio of the simulation series are indications of the load imposed by the need to interpret each set of values. For a particular case, Num subjects were more likely to run the simulation several times, seemingly in an attempt to gain an appreciation of the form of the collision. The LED subjects, on the other hand, can quickly get an appreciation of the form of a collision by looking at the 1DP diagram. The direction of the arrows and their lengths makes the form of the collision available by means of a quick perceptual inferences. Thus, it appears that LEDs subjects may have been comparing successive collisions, rather than individual cases, given their greater proportion of value changes and ratio of value change series.

The explanation of the difference between the LED group and the Num group goes deeper than the fact that the LED subjects could more quickly examine a greater number of collisions. Remember that the measures of value changes and runs of the simulations are relative to the total number of operations executed by each subject; thus the rate at which LED subjects examined different collisions was greater than the Num subjects. This, and the fact that the LED group examined a greater number of different collisions (higher configuration distribution measures), implies that the LEDs seem to make it easier for the subjects to consider different forms of collisions. In the 1DP diagram and the V-V graph, the different collision are distinct patterns or shapes, so considering different forms or configurations requires relatively less cognitive effort than attempting to distinguish different cases from sets of numerical values. Note that the availability of a table of previous values did not appear to help the Num group. To some extent, the LEDs appear to make more of the space of configurations of collision accessible to the system users.

General Discussion

Some general implications for instruction and learning in technical domains follows from the findings of this study. Providing computer based diagrams that make the underlying relations of a domain more accessible to problem solvers may be an effective way to facilitate the learning of qualitative relations. LEDs make the underlying laws of a domain more accessible by capturing them in the internal structure of diagrams and by having each diagram represent one instance of the phenomenon. This means that the learner is less burdened with the task of interpreting individual cases, so is freer to examine successive cases or to consider extreme and special cases. Such cases usually correspond to distinctive and unexpected patterns in LEDs, which will challenge and help refine the learners understanding of the rules governing the form of the LEDs. Hence, there seems to be a double benefit in making the relations readily accessible.

The findings of this experiment are consistent with the previous small scale experiment on ReMIS-CL (Cheng, 1994, 1995). In that experiment the subjects who successfully used the LEDs in the post-test were the ones who had obtained the most complete understanding of the constraints of the LEDs. They achieved this by exploring more of the space of possible configurations of the LEDs. Both experiments suggest the importance of examining a wide range of cases during learning with LEDs, but this same approach may be effective with other representations. By exploring more of the space of possible configurations of the phenomenon, the learner will see a greater variety of the correct forms of expressions in a representation and thus have greater variety of distinct cases over which to induce the relations underlying the phenomenon. Investigating extreme cases is more likely to provide useful information for establishing the boundary conditions of a law than merely examining normal or typical cases.

The LED approach is now briefly contrasted with some examples of other research on computer based systems for physics learning. Whitelock *et al.* (1993) have also studied learning about collisions and have shown that faulty causal models can be challenged with a computer simulation. This is consistent with White's (1993) findings. Although not in the same domain, White's approach is interesting because it provides learners with correct *intermediate causal models* (ICMs), in addition to animated computer simulations. ICMs act as bridge for the conceptual gulf between abstract general laws and observations of phenomena. Children have successfully learnt about Newtonian dynamics with ICMs, achieving a level of understanding matching students beginning undergraduate science programs.

LEDs may also be considered as representations at an intermediate level of abstraction. However, LEDs capture the formal relations defined by the laws of a domain in a constraint based manner, rather than portraying causation.

Ploetzner *et al.*'s (1990) DiBi system is an intelligent tutoring system for elastic collisions that aims to support and guide students in their construction of sound domain representations. In contrast, ReMIS-CL is a discovery learning environment that has no built in intelligence or student model, but it explicitly provides correct domain representations in the form of LEDs.

The present study has demonstrated the effectiveness of LEDs for learning about some qualitative relations and extreme cases in the domain of elastic collisions. It has also provided a better understanding of some of the processes involved in effective learning with LEDs. However, many issues are raised by this and the research mentioned above. Will LEDs be useful to students attempting to develop conceptual understanding of a domain? Can LEDs be effective for students who have less experience in physics and who are much younger? How important is it to have two complementary LEDs, such as the 1DP diagram and the V-V graph, in a single system? Ongoing research is attempting to address these questions.

Acknowledgements

The research was supported by the U.K. Economic and Social Research Council. I am grateful to the Physics Department at the University of Nottingham, especially John Owens-Bradley and Chris Mellor, for helping me to set up this experiment. Thanks must also go to members of the ESRC Centre for their assistance in my pursuit of this work, and especially to Fernand Gobet for his valuable comments on a draft on this paper.

References

Cheng, P. C.-H. (1994). An empirical investigation of law encoding diagrams for instruction. In *Proceedings of the 16th Annual Conference of the Cognitive Science Society*. (pp. 171-176). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cheng, P. C.-H. (1995). Law encoding diagrams for instructional systems. *Journal of Artificial Intelligence in Education*, 6(4).

Cheng, P. C.-H. (in press). Scientific discovery with law encoding diagrams. *Creativity Research Journal.*

Cheng, P. C.-H., & Simon, H. A. (1992). The right representation for discovery: Finding the conservation of momentum. In D. Sleeman & P. Edwards (Eds.), *Machine Learning: Proceedings of the Ninth International Conference (ML92)* (pp. 62-71). San Mateo, CA: Morgan Kaufmann.

Ploetzner, R., Spada, H., Stumpf, M., & Opwis, K. (1990). Learning qualitative and quantitative reasoning in a microworld for elastic impacts. *European Journal of Psychology of Education*, V(4), 501-516.

White, B. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, 10(1), 1-100.

Whitelock, D., Taylor, J., O'Shea, T., Scanlon, E., Sellman, R., Clark, P., & O'malley, C. (1993). Challenging models of elastic collisions with a computer simulation. *Computers in Education*, 20(1), 1-9.