Interactive lecture demonstrations: a tool for exploring and enhancing conceptual change

Rachel Zimrot and Guy Ashkenazi*

Department of Science Teaching, The Hebrew University of Jerusalem, Israel
e-mail: guy@fh.huji.ac.il

Received 6 November 2006, accepted 13 March 2007

Abstract: Interactive Lecture Demonstrations (ILD) are a student centered teaching method, in which students are asked to predict the outcome of an experiment, observe the outcome, and discuss it with respect to their former expectations. The demonstrations are designed to contradict students' known misconceptions, generate cognitive conflict and dissatisfaction with the existing conception, and promote a process of conceptual change. An ILD based course was used to explore the effect of cognitive conflict on the conceptual change process, and the role of student interactivity in this process. Three major levels of conceptual change were identified: high – students who remember the outcome of the demonstration, and explain it using the consensus model; medium – students who can recall the outcome, are dissatisfied with their alternative model, but do not switch to the consensus model; and low – no meaningful recollection of the outcome, and no change in the alternative model. A multiple-choice test based on the lecture demonstration was given to two groups, one of which only observed the demonstrations, without predicting and discussing. We found a significant difference between the groups, with an obvious drop in students' ability to recall the outcome of the demonstrations in the non-interactive group. [*Chem. Educ. Res. Pract.*, 2007, **8** (2), 197-211]

Keywords: Interactive Lecture Demonstrations (ILD), conceptual change, cognitive conflict, mental models, levels of conceptual change

Introduction

Science is a human endeavor to interpret natural phenomena in an efficient and consistent way, which allows accurate predictions and coherent explanations. Students often interpret natural phenomena in naïve ways, which may differ radically from what modern science considers as efficient or consistent (Driver et al., 1985). Such personally constructed interpretations have proved difficult to change by traditional instruction. One important goal of research in science teaching is to construct methods guiding learners to think in terms of efficient and consistent scientific concepts. This means helping individuals to transform their *mental models* (internal representations of objects, events and processes individuals construct in order to predict and explain phenomena) into the *consensus models* (the expressed representations used by the scientific community for the same purpose) (Gobert and Buckley, 2000). This is a process of *conceptual change*. The change from naïve mental models to consensus models is seldom straightforward, and many times it goes through intermediate stages which combine parts of both models – a *hybrid model* (Galili et al., 1993). In this work, we refer to both naïve and hybrid models as *alternative models*, to distinguish them from the scientifically accepted consensus models.

A major difference between alternative models and consensus models is in the range of their validity. While alternative models are usually only consistent with the limited experience

This journal is © The Royal Society of Chemistry

base of the individual, the consensus models are based on the collectively accumulated experience of generations of scientists. The efficiency of consensus models arises from their consistent capability to predict and explain a vast range of phenomena using a small set of assumptions and rules. Therefore, to facilitate the process of conceptual change, students should become familiar with a larger set of natural phenomena than they normally encounter in everyday life. For this purpose, we devised an extensive set of lecture demonstrations, which introduce students to phenomena that cannot be easily predicted or explained by their naïve personal models.

Our approach was partly inspired by the work of Posner and his colleagues (Posner et al., 1982; Strike and Posner, 1992). Their theory of conceptual change is based on the historical development of science. They argued that students will be willing to change their alternative models only if they are dissatisfied with them, and the proposed model proves to be more fruitful. Therefore, students should be confronted with discrepant events and anomalous data their existing models cannot account for, in much the same way that new discoveries in science challenged existing theories. Concurrently, the students should be exposed to a consensus model which is able to account for the new data. In light of this theory, methods for promoting conceptual change have focused on establishing conditions where the student's existing conception can be made explicit, and then is directly challenged to create a state of conflict. However, Posner et al. (1982) admitted that while such a *cognitive conflict* is a necessary condition for conceptual change, they do not expect its mere introduction to be sufficient to induce the change, nor to make the process straightforward.

More recent studies of the effect of cognitive conflict on promoting conceptual change have corroborated this expectation (Limon, 2001; Mason, 2001). Instead of directly addressing the anomalous data and modifying their existing conceptions, students exhibit a diverse range of behaviors in response to being confronted with discrepant events (Chinn and Brewer, 1993, 1998; Tirosh et al., 1998; Shepardson and Moje, 1999; Mason, 2001; Kang et al., 2004). For example, Chinn and Brewer (1998) describe a progression in the ways in which undergraduate students respond to reported anomalous data: (1) Ignoring – the outcome simply goes unnoticed; (2) Rejection – denying the validity of the discrepant event and explaining the reasons for rejection; (3) Uncertainty – not sure if data is believable; (4) Exclusion – considering the discrepant event as irrelevant; (5) Abeyance – expecting that their existing model might be able to explain the anomalous data in the future; (6) Reinterpretation – making peripheral changes to the data so the existing model can fit it; (7) Peripheral theory change – Accepting the data and making peripheral changes to the existing model; and (8) Theory change – Accepting the data and fundamentally changing their model to accommodate the new data. In all the above scenarios save the last, the students' confrontation with anomalous data did not result in conceptual change. In many of the scenarios, the students found a way to avoid the conflict, and reconcile the anomalous data without having to consider their existing model.

In order to encourage students to actively employ their existing models in the context of lecture demonstrations, and thus make conceptual change more likely, we followed the Interactive Lecture Demonstration (ILD) teaching strategy (Sokoloff and Thornton, 1997). In this method, the students are asked to predict the outcome of a demonstration and write down their prediction and an explanation, and therefore commit to an explicit model. Peer discussion follows, with the students discussing their predictions in small groups – again, they have to address their existing models explicitly. The instructor engages the class, soliciting predictions and highlighting common predictions. The demonstration is then performed, and the instructor discusses the results of the demonstration in view of the students' previous predictions, emphasizing the conflict and providing the consensus way to resolve it.

This journal is © The Royal Society of Chemistry

In addition to enhancing students' commitment and explicit involvement in the conceptual change process, the highly interactive nature of the teaching method allowed us to investigate the different ways in which the students were involved in this process. To this end, we conducted a combined qualitative and quantitative study regarding the effect of interactive lecture demonstrations on students' conceptual change process. Our research questions were:

1. What are the possible effects of the cognitive conflict, triggered by the lecture demonstrations, on students' conceptual change process?
2. What is the contribution of the interactive part of the ILD method, as compared with passive lecture demonstrations?

Pedagogical Context

The study was conducted over a period of three years, starting in the fall semester of 2001, while the second author was teaching the course 'General Chemistry for Biology Majors'. This is a one-semester freshman course, given each fall, with an average enrollment of 200 students.

Content

The fall semester is 14 weeks long. One week was devoted to each of the following topics:

1. Chemical change and conservation of mass.
2. Particulate nature of matter.
3. Electronic structure of metallic, ionic and molecular substances.
4. Quantities in chemistry: mass, volume and number.
5. Electrolytic decomposition and reactions in aqueous solution.
6. Thermal energy, motion and temperature.
7. Interconversion of chemical and thermal energy.
8. Theories of bonding and molecular structure.
9. Intermolecular interactions and physical properties.
10. The ideal gas law and the kinetic molecular theory.
11. Collision theory and the rate of chemical reactions.
12. The dynamic nature of microscopic equilibrium.
13. Acid-base equilibrium (2 weeks).

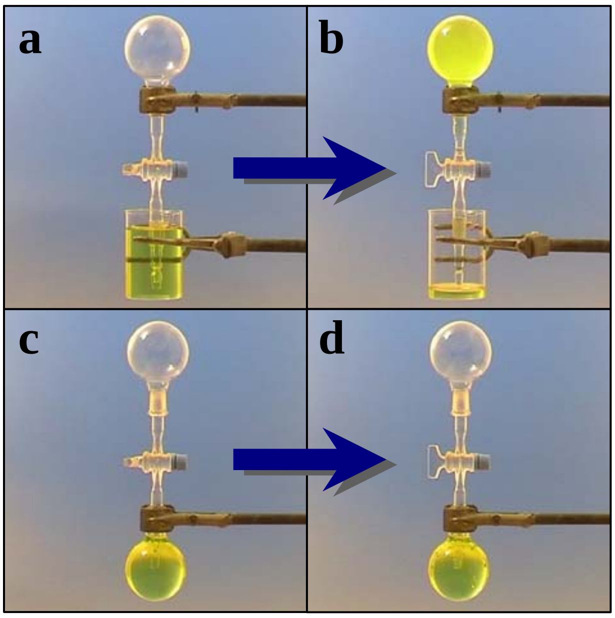
Each week, the class met for two 90 minute lecture periods, plus one 90 minute small group recitation session.

Teaching Method

One lecture period each week was devoted to the discussion of lecture demonstrations. For each topic (except for topic 8), we developed a sequence of 3-5 demonstrations, all of which can be explained using the same consensus model. Some of the demonstrations were designed with the expectation that most students will predict the outcome correctly, even if they just use common sense gained from everyday experiences. However, at least one demonstration in each sequence was designed to contradict predictions based on common alternative models, as known from existing misconception literature (Pfundt and Duit, 2000) and our own pilot study (to be described under the Methodology section). Figure 1 shows an example of a common sense experience vs. the *conflict demonstration* in topic 2, designed to counter models which consider vacuum as exerting a pulling force (Nussbaum, 1985). The demonstrations were projected on a large screen, using a video camera connected to the classroom's projection system, so that all students could see them clearly.

This journal is © The Royal Society of Chemistry

Figure 1. Common sense experience vs. conflict demonstration.



(a) The demonstration system is composed of an evacuated glass bulb (top), connected through a closed valve to an open glass containing water with green food coloring. (b) When the valve is open, the water is pushed by the pressure of the atmosphere into the evacuated bulb, creating the common sense feeling of suction, as if the vacuum was 'pulling' the water up. (c) The experiment is repeated, this time with an air-tight closed glass bulb at the bottom. (d) When the valve is opened, the water stays at the bottom – the vacuum by itself has no 'pulling' force, and the rigid wall of the container keeps the atmospheric pressure from pushing the liquid up. For a video of the demonstrations, visit http://www.fh.huji.ac.il/~guy/links/CERP2007a.html .

The structure of the lesson for each demonstration was:

1. The instructor shows the class the experimental setup, and explains what he is about to do.
2. The students select one of the pre-determined possible outcomes, and write down an explanation for their prediction on a Prediction Sheet (Figure 2), which is collected at the end of class.
3. The students discuss their predictions with their neighbors for 3-5 minutes, and then vote for one of the possible outcomes.
4. The instructor displays the distribution of predictions, and solicits explanations from the class for the different opinions.
5. The instructor performs the demonstration, noting which of the predictions proved correct. Students record the results on a Results Sheet, which they keep.
6. The instructor projects a PowerPoint slide that offers different explanations for the observed experimental result, only one of which uses the consensus model (Figure 3). The other explanations are based on known students' alternative models. The students discuss the alternatives among themselves and vote for the best explanation.
7. The instructor discusses the outcome of the experiment using the consensus model, while explicitly referencing students' alternative models and noting their deficiencies and limitations.

The prediction and discussion prior to the demonstration help students to construct an explanation based on their personal models. The outcome tests the validity of these models and demonstrates their limitations (generates a cognitive conflict). The discussion of alternative explanations, after seeing the outcome, illustrates the wider domain of validity of the consensus model over the alternative models, and therefore establishing its fruitfulness.

This journal is © The Royal Society of Chemistry

Figure 2. Part of the prediction sheet for topic 2.*

Demo 3

The nozzle of an evacuated bulb is inserted into a flask filled with water.

Predict the state of the system after the valve is opened.

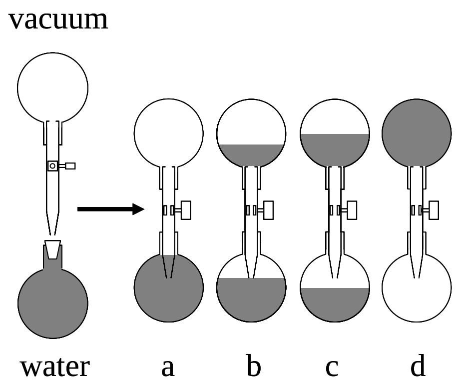
Provide a short explanation:

____

____

____

____



* The pre-determined distracters for the predictions (a-d above) were taken from the pilot study. The complete sequence included four demonstrations. The first is depicted in Figure 1a and 1b; the second used the same system as above with NO_2 (a brown gas) instead of water; the third is depicted above; and the fourth used NO_2 + air (instead of vacuum). For videos of all the demonstrations, visit http://www.fh.huji.ac.il/~guy/links/CERP2007a.html .

Figure 3. A follow-up discussion question for the conflict demonstration.

Which of the following statements best explains the observed phenomena (the liquid stays down)?*

1. If the liquid went up to fill the vacuum on top, vacuum would be formed at the bottom.
2. The particles in the liquid are attracted to each other, and to the earth.
3. The negative pressure by itself is not strong enough to overcome the weight of the water.
4. The liquid has the property of keeping its volume, and filling the bottom of the flask.

* Explanation 1 is based on the model that vacuum causes water to move because matter has an aspiration to fill empty space. Explanation 3 is based on the model that vacuum is a negative pressure that pulls on the water. Explanation 4 is correct, but is limited to liquids. Explanation 2 uses the consensus model which analyzes the system in terms of particles and forces between them.

Methodology

In educational research, one must take into account the complexity of human thinking and its influence on the process of learning. Different individuals often act differently under the same circumstances, and the same individual may act differently under different circumstances. Before trying to assess the impact of a new teaching method at the population level, one must first get acquainted with the possible different ways in which each individual

This journal is © The Royal Society of Chemistry

can be affected. We therefore employed a combined qualitative-quantitative methodology in this study.

Pilot Study

In fall 2001 we ran the course for the first time, in order to fine tune the teaching method, and construct a research based questionnaire for the quantitative study. The teaching method was similar to the one described above, with two changes: there were no pre-determined possible outcomes in step 2, and no pre-determined alternative explanations in step 6. Instead of selecting a prediction from a list of options, the students were asked to draw their own prediction on the prediction sheet, before writing their explanations. The prediction sheets were collected at the end of the lesson, and students were given credit for handing them in, regardless of the correctness of their responses.

We used the prediction sheets to learn about students' alternative models in the context of the demonstrations. The collected data was analyzed according to the facets-schemes method (Galili and Hazan, 2000). Initially, responses which seemingly presented the same meaning were grouped together, even if they were expressed in somewhat different wording. Then, we identified representative categories of explanatory patterns or strategies employed by the students in addressing particular situations. These are the observable facets of knowledge. At the last step, we grouped facets of alternative knowledge elicited in different experimental contexts around a smaller number of less specific explanatory models. These models represent the tacit underlying schemes of knowledge.

Following the analysis, we completed the lesson plans for the subsequent years. We chose pre-determined predictions for step 2, and alternative explanations for step 6.

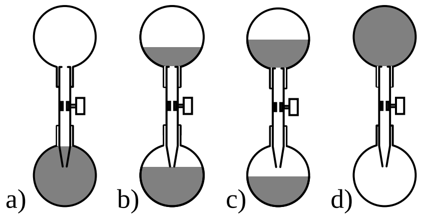
Next, we constructed a conceptual test based on the facets-schemes analysis. In every topic, we chose a single demonstration which proved to be the hardest to predict correctly, and constructed two multiple-choice questions addressing the specific demonstration. The first question in each pair is a prediction question, for which we chose the most common students' predictions as distracters. The second question asks the student to select the scientific explanation out of five options, four of which correspond to the most common alternative models. Figure 4 shows sample questions from two topics.

The questionnaire was aimed to measure whether students accept the anomalous data provided by the conflict demonstration, and whether they can recognize explanations based on the consensus models taught in the course. The completed questionnaire was examined for correctness and intelligibility by one other chemistry professor and two high-school chemistry teachers. Since the questions were taken directly from the teaching materials, content validity is inherent.

This journal is © The Royal Society of Chemistry

Figure 4. Part of the conceptual test, for topics 2 and 11.

3. The nozzle of an evacuated bulb is inserted into a closed flask filled with water, and the valve is opened. How will the system look after one minute?



4. What is the accepted scientific explanation for this phenomenon?

a) The liquid flows from high pressure to low pressure in order to equalize pressures.

b) The liquid flows up in order to fill the empty space.

c) There is no force pulling upward.

d) Gravitation pulls the liquid down and the vacuum pulls it up.

e) The liquid conserves its volume, therefore the system cannot act to fill the vacuum.

19. 40 mL of 3.0 M HCl(*aq*) react with excess limestone in an open flask, releasing carbon dioxide gas according to the equation:

$$CaCO_3(s) + 2HCl(aq) \rightarrow CaCl_2(aq) + H_2O(aq) + CO_2(g)$$

The graph of the change in weight of the reaction flask in the first minute is linear, with a slope of -0.3 g/min. The experiment is repeated with 20 mL of 6.0 M HCl(*aq*). In both cases the acid covers all of the stone. What will the slope of the graph of the change in the system's weight be, in the first minute?

a) Steeper than -0.3 g/min.

b) -0.3 g/min.

c) Less steep than -0.3 g/min.

20. What is accepted scientific explanation for this phenomenon?

a) The rate of collisions between the reactants' particles is greater.

b) The number of moles of acid is equal.

c) Larger volume means a larger surface area.

d) There is less acid in the second experiment.

e) The stronger acid can overcome the activation barrier more easily.

Qualitative Study

In this part of the study, we characterized the levels of conceptual change attained by the students. In other words – do they think about phenomena in terms of consensus models, or do they still use their naïve mental models?

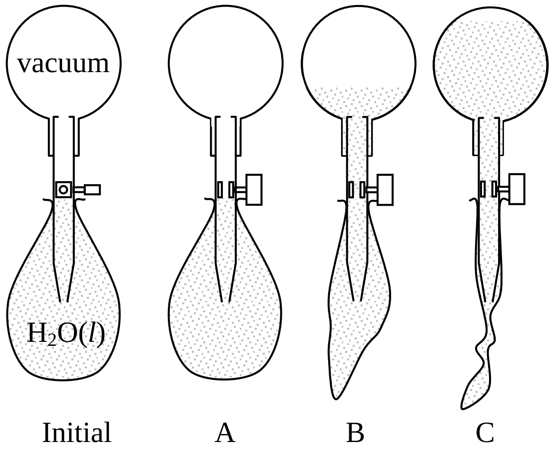
This study was conducted during the fall 2002 semester. From all the students who fully participated in the class in the first 4 weeks (filling all prediction sheets completely, and solving all homework problems), twelve were picked for personal semi-structured interviews. Each student was interviewed 3 times throughout the semester, at week 4-5, week 8-9 and week 14 (end of the semester). Each interview was approximately an hour long, and included questions about learning in general, and specific content questions. There were two types of content questions: recall and near transfer. Recall questions addressed the demonstrations that were presented in class: what happened, how they explained the outcome in the prediction sheet, how would they explain it now (during the interview), and (in the relevant cases) why did they change their explanations. Near transfer questions concerned a similar system, but in a slightly different context. Figure 5 shows an example of a near transfer question (compare to Figure 2).

When analyzing the interviews, we compared the students' models that emerged from the interviews to the models that emerged from their prediction sheets. The comparison focused on three key questions, adapted from the work of Chinn and Brewer (1998): (1) Can the

This journal is © The Royal Society of Chemistry

student recall or reconstruct in his mind the conflicting outcome of the demonstration? (2) Does the student offer an explanation to the outcome of the demonstration? (3) Does the student accept the consensus model?

Figure 5. A near-transfer interview question.*



* This system is similar to the one depicted in figure 1c, except that the water is held in a flexible plastic bag, instead of a rigid glass bulb. The student is asked to predict what will happen to the system when the valve is opened (1) on the Earth, and (2) on the moon.

Quantitative Study

This part of the study examined the contribution of the interactive part of ILD to the overall effect of lecture demonstrations on students' conceptual levels. We used an experimental comparative design. Because we consider ILD to be the norm in our course, the experimental 'treatment' was the discontinuation of the interactive part half way through the semester. This was done in fall 2002. After week 8 (that didn't have any demonstrations associated with it), the instructor stopped using the interactive components (steps 2, 3, 4 and 6 of the teaching method) in his lectures. The instructor still carried out all the demonstrations (steps 1 and 5), and discussed students' alternative models after showing the result of the demonstration (step 7). However, these models were not solicited from the students in class, but merely presented to them as experience gained from the pilot study ("*When I taught this topic last year, many students thought that…*"). Skipping the prediction and discussion part of the demonstration freed up a lot of lecture time. Since the same time was still allotted for each topic, the lectures in the second part of the course included more detailed explanations by the instructor regarding the consensus model.

The control group, in fall 2003, continued to learn using ILDs throughout the entire semester. The two groups were comparable in terms of academic achievement. At the end of the study, we obtained the group grade average of participating students in six other courses (in biology, physics and math), which showed no significant difference (t = 1.44, p > 0.1) between the treatment group (85.2 ± 11.5, n = 115) and the control (83.0 ± 13.0, n = 141).

In both groups, we administered the multiple-choice conceptual test on the last day of the course. The students weren't informed about the test in advance, and therefore the test reflects the knowledge and understanding they constructed during the course, before they started preparing for the final exam.

This journal is © The Royal Society of Chemistry

Results

Qualitative Study

From the interviews, we have identified three major levels of conceptual change. Each level is associated with a different degree of impact that the cognitive conflict had on the student's prior model. To promote clarity, we limit the presented evidence to students' citations that pertain to topic 2, which was described in detail in the previous sections. This topic had the largest percent of wrong in-class predictions – only 1/3 of the students, each year, predicted the conflict demonstration correctly (option a in Figure 2), while the other 2/3 expected at least part of the water to go up into the vacuum (nearly equally split between options b and c in Figure 2). Similar levels of conceptual change were identified across all topics.

High Level

At this level, the students recalled what they saw in class, they recalled their previous explanation, and they changed their explanation in the interview to one which conforms with the consensus model. In most cases, the students declared that the conflict demonstration helped them to sort out their conceptions and to change them. For example:

> Q: In demo 3 [Figures 1c, 1d], you now say that you remembered that the water stayed down, but you predicted something else, as you can see in your sheet. What did you feel when you saw a different outcome?
>
> A: It changed my conception. It changed my conception [repeating, for emphasis]. Like, my previous conception was from that I had no knowledge of chemistry, and I didn't deal with such things. And my conception was that the vacuum is really a sort of suction force, and because of this I thought the water will go up. […] There will be here the same vacuum as here [points at the top and the bottom bulbs in option 'c' in Figure 2], because it is a kind of a suction force, and will result in an equalization of pressures. But after we, like, saw – I understood why it is not like that […]
>
> Q: When you saw it, did you change your conception at the same moment?
>
> A: No. I was very surprised when I saw it. I'm speaking about during the course, after the instructor explained it to us, and after I read about these topics […] And especially after I saw it, that it really doesn't happen. Because of this I think these demonstrations are very important. Because if you just came and told me, like tell me in class that someone did this experiment, show me these drawings and tell me 'this is what happened', I probably would not remember it now. But I remember it because I remember it surprised me, and I remember I pondered over it. Like I didn't understand why it is so. It is a process.

There was some variability within this broad level. Some of the students that successfully used the consensus model to explain the lecture demonstration, did not succeed in predicting or explaining the answer to the near-transfer question by themselves. However, these students were able to answer correctly after being prompted by the interviewer to go back to the prediction sheet and look for a similar situation they have already seen in class.

Medium Level

At this level the students recalled what they saw in class, they recalled their previous explanation, and they were aware that it was inadequate to account for the outcome of the conflict demonstration. However, the model they used to explain this outcome in the interview was not the consensus model, but a different alternative model. They were often unsure about their new model, and used it hesitatingly. For example:

> [While answering the near-transfer question in topic 2 – Figure 5] Maybe the [atmospheric] pressure can influence outside of the bag, and it will be half-and-half […] because then there will be equal pressures. There is gravity that pulls downward, and on the other hand it is pulled [up]

This journal is © The Royal Society of Chemistry

because of the vacuum, then it probably equalizes there […] I learned in class that vacuum is not a force, but somewhere… [hesitates] it is still like… but the vacuum pulls! I don't remember what caused the vacuum to pull – this is the principle I'm trying to recall. I know it is not a force, but there is something in it that causes a force.
Q: Let's get back to the lesson. In demo 3 [Figures 1c, 1d], the instructor opened the valve. Do you remember what happened?
A: It seems to me now that the water didn't go up. I don't remember. [Asks for verification] Yes? There is no atmospheric pressure to push them up. There is just the vacuum that I don't remember what it does. […]
Q: You wrote that the vacuum will cause it to go up. Why?
A: Back then I was sure that vacuum is a force, and when you have vacuum it pulls. I didn't know then about the force of atmospheric pressure, about all these things. […]
Q: When you see things that don't behave as you expected…
A: It amazed me. I didn't have the smallest bit of idea why it didn't go up. I tried to think why it happened, but no idea came to my mind.
Q: When you think of something and it comes out the opposite, does it frustrate you or do you want to know why?
A: I want to know why, because it annoys me that I'm not right!

Again, there is some variation in this category. Some students just switched from one naïve model to another. Others, like the one in the example, made an effort to incorporate the consensus model into their existing scheme. While this student struggled to make sense of the new model, she did not abandon her previous model. This resulted in a hybrid model, in which the pulling force of vacuum complements the pushing force of the atmospheres.

Low Level

Students that belong to this group did not remember the outcome of the demonstration. Without a vivid memory of the demonstration, they could only rely on their alternative models when they tried to recollect what happened in lecture. This resulted in students giving incorrect predictions to experiments they had actually observed in the past. For example:

Q: Can you explain what you saw in demo 1 [the common sense experience in Figures 1a, 1b]?
A: I remember that what I was thinking wasn't correct [laughs in embarrassment], but I still remember what I thought then.
Q: What did you think then?
A: That there is some kind of aspiration of matter to spread out into the vacuum. That is, it will flow in the direction of the vacuum, and fill the bulb. But I think this is not the correct explanation. […] It is like drinking with a straw, where the vacuum is equivalent to a pressure acting on the liquid, pulling it up – it [the liquid] seeks to fill it [the vacuum]. […] It is not really a force – the vacuum has no force, but because matter seeks to fill it, so to speak, then the pressures' differences are those that push it up.
Q: Do you remember what happened in demo 3 [Figures 1c, 1d]?
A: [Asks herself hesitantly] Could it be 'd' [in Figure 2]? [Laughs in embarrassment] I don't remember. [Suddenly sounds more sure] OK – there is a vacuum on top, so it would equilibrate, and it could be c as well. It is actually a balance between the two bulbs, something like that, because it is a closed system, and the pressures between the bulbs want to be balanced and will reach some state of equilibrium.
Q: This is what you wrote in the prediction sheet. Let me tell you what actually happened – the water didn't go up.
A: OK… [sounds doubtful, gives a puzzled look]
Q: When I remind you now, can you recollect why it didn't go up?
A: I really don't remember. [Asking herself] Why didn't it go up? Maybe because it is a liquid, contrary to a gas, where the gas spread out in the second demo and filled the container, and the water didn't.

This journal is © The Royal Society of Chemistry

This is an example of a student who cannot let go of her prior scheme that matter flows purposefully in order to fill the vacuum. Even though she's aware that her explanation is incorrect, she still used the same argument when answering the near-transfer question. Her inability to explain the conflict demo in class has led her to ultimately ignore this data, and she couldn't remember the fact that it contradicted her expectation. When re-confronted with the anomalous data, she chose to exclude it, by suggesting that while her model works well for gases, liquids might behave differently.

Other variations of low level responses included students who just declared they did not understand the subject matter, and a few cases of students who actually believed that they saw what they expected to see.

To summarize, the three levels differ in their ability to recall and explain the phenomena they saw, and this ability is connected to the degree to which they changed their prior models. Students at the low level of conceptual change cannot recall the outcome of the conflict demonstrations. They have not been affected by the conflict, and did not undergo any conceptual change. Students at the medium level of conceptual change can recall the outcome of the conflict demonstration, but cannot give a scientifically accepted explanation. They have been affected by the conflict, but their conceptual change is limited, because they could not give up their alternative models. At the high conceptual change level, students can both recall and give a scientifically accepted explanation. They have completed the process, as they started out with an alternative model, and can now use the consensus model.

Quantitative Study

In light of our findings in the qualitative research, we classified students' responses to each pair of questions on the conceptual test (Figure 4) into similar categories. A student's understanding of a specific topic was classified as *high conceptual level* only if both the prediction (first question) and explanation (second question) were correct; as *medium conceptual level* if the prediction was correct but the explanation was incorrect; and as *low conceptual level* if the prediction was incorrect, regardless of the correctness of the explanation.

Since the goal of science teaching is for students to achieve a high conceptual level, we calculated the students' *conceptual score* by counting the number of topics in which they attained a high conceptual level (both explanation and prediction are correct). Because we changed the teaching method for the treatment group partway through the semester, we calculated two conceptual scores for each student – a score for Part I of the course (topics 1-7, maximum score = 7), and a score for Part II (topics 9-13, maximum score = 5). By looking primarily at the high conceptual level we also minimize the effect of random guessing, because the chances of correctly guessing both the prediction *and* the explanation are between 1:10 and 1:20.

To determine the effect of the experimental change in teaching method, we need to compare the two groups' conceptual score in Part II, in which passive lecture demonstrations were used with the treatment group. Because the groups are not necessarily equivalent, we need to control for any disparity that might affect the difference in this conceptual score between the two groups. We expected, and found, a statistically significant correlation between the students' conceptual scores on part I and on part II, with $R^2 = 0.26$, and $p < 0.001$ (this means that 26% of the variance in the score on part II can be explained by the variance in the score in part I). It is safe to assume that the students' scores on Part I were not significantly affected by the different treatments they received in Part II. We can therefore use the conceptual scores in Part I as a covariate in ANCOVA, and thus account for possible disparity between the groups. The conceptual score in part II is the dependent variable, and interactivity in the teaching method is the independent variable. The results of the ANCOVA

This journal is © The Royal Society of Chemistry

are summarized in Table 1, and show a statistically significant difference in conceptual score between the groups (*p* < *0.001*). There is a medium positive effect (*d* = *0.48*) of the interactive method over passive lecture demonstrations.

Table 1. ANCOVA results for the effect* of interactivity on students' conceptual achievement score (0-5) in Part II of the course.

Adjusted Means ± SD			
Interactive LD	Passive LD	*p*	*d*
2.55 ± 1.16	2.02 ± 1.10	0.0002	0.48

* *d* is Cohen's effect size. 0.48 is considered a medium effect.

To better understand the meaning of this difference, it is instructive to see how the distribution of conceptual levels changes when interactivity is discontinued, as shown in Table 2. In the fully interactive course (control group), only ~30% of students' responses are at a low conceptual level, and more than 50% of the responses show a high conceptual level. Similar numbers appear in the first part of the treatment, which was also taught with ILD. However, when interactivity is discontinued in the second part of the treatment course, this situation reverses – as many as 51.8% of the responses are at a low conceptual level, and only 38.3% are at the high conceptual level.

Table 2. Percent distribution of conceptual levels, in each part of the two runs.

	Interactive Throughout		Interactivity Discontinued	
Conceptual Level	Part I (ILD)	Part II (ILD)	Part I (ILD)	Part II (LD)
High	58.6	52.8	51.9	38.3
Medium	15.1	12.8	15.9	9.9
Low	26.3	34.4	32.2	51.8

LD = passive lecture demonstrations.

It is also interesting to note that in all four cases a large majority of the students' responses is either at a high or a low conceptual level, and only a small part (10%-15%) is at a medium level.

Discussion

Answer to first research question

The first research question in this study was: "*What are the possible effects of the cognitive conflict, triggered by the lecture demonstrations, on students' conceptual change process?*" Our results show that the effects are varied, and conceptual change can be an evolutionary, rather than revolutionary, process. When confronted with a discrepant event, students may respond in different ways, and attain different levels of conceptual change, that are in overall agreement with the *taxonomy of responses to anomalous data in science* proposed by Chin and Brewer (1998). However, there are a few points in which this taxonomy falls short of describing important features that are relevant to teaching with lecture demonstrations, which we would like to emphasize.

The high conceptual change level matches *theory change* in the taxonomy, in which individuals abandon their former belief in favor of a new one. We noticed that the term

This journal is © The Royal Society of Chemistry

'abandon' might be too strong, as some students who attained this level for the lecture demonstration still used their previous model in other contexts (the near-transfer question). As novices, they probably still lack the ability to recognize novel situations in which the newly acquired model would be applicable and fruitful. Nevertheless, after accepting the new model in one context, it was fairly easy for them to apply it, upon prompting, in other contexts. When students attained a high level of conceptual change in a specific topic, they often associated this achievement with their positive experience in resolving the cognitive conflict induced by the lecture demonstration.

The medium conceptual change level matches *peripheral theory change*, in which individuals make minor changes to their theories without giving up the core components. These students were aware of the conflict between their existing model and the outcome, tried to change their explanation to accommodate the outcome, but still did not accept the consensus model. This was because the students found it hard to relinquish their prior beliefs, or because of poor understanding of the consensus model. This process was accompanied by a feeling of struggle between the two models and evident discomfort, which were a direct result of the unresolved cognitive conflict. The quantitative data shows that this level is the least populated – only 10%-15% of the answers to the conceptual test fall in this category, in which the student remembers the outcome, but fails to explain it using the consensus model. This might indicate that students do not stay in such a state of discomfort for a long period of time. However, we can't tell if this means that these students resolved their conflict by embracing the consensus model, or by receding to a lower level and rejecting the anomalous data.

The most surprising to us was the large percentage of responses which fell into the low level of conceptual change. At this level, students fail to recall the outcome of a demonstration that was performed in front of their eyes, and contradicted their expectations. This matches several of the low level responses in the taxonomy – *ignoring*, *rejection*, and *exclusion*, in all of which the individual does not accept the data. The students at this level hold so strongly to their existing models, that they fail to experience the cognitive conflict. To use a cliché, instead of saying "*I had to see it to believe it*", such students should say "*I had to believe it to see it*". They are probably aware of the conflict at the time of the demonstration, but their inability to formulate a meaningful explanation for what they saw leads to rapid fading of any memory of the discrepant event. It seems that the cognitive conflict method fails for those students who need it the most – students who have a difficulty to change their conceptions.

Answer to second research question

The second research question was "*what is the contribution of the interactive part of the ILD method, as compared with passive lecture demonstrations?*" Our results show that the interactive component of the lecture demonstration plays an important role in promoting students' conceptual change. Its discontinuation at the middle of the semester resulted in a statistically significant drop in students' conceptual score. The rate of high conceptual level answers decreased, and the rate of low level conceptual answers increased. Without the interactive component, more than half of the students saw a demonstration which should have contradicted their expectations, but had no meaningful recollection of this as a discrepant event. This result is in agreement with the work of Crouch et al. (2004) who studied different modes of lecture demonstrations in physics. They found that just observing a demonstration gave students little advantage over not seeing a demonstration at all, especially in terms of their ability to explain the outcome of the demonstration. Asking the students to predict the outcome of the demonstration before it was carried out had a significant impact on their ability to repeat the prediction correctly and give an acceptable explanation at the end of the course.

This journal is © The Royal Society of Chemistry

The contribution of the interactive component of the lecture demonstrations can be attributed to two factors. The first is that the prediction and discussion steps make the students explicitly aware of their existing models. This increases their awareness to the conflict, and their need to resolve it. It is harder to ignore an unexpected result after you have committed to it in writing, and explained it to your friends. The second is that ILD can be classified as an Interactive Engagement teaching method – designed to promote conceptual understanding through interactive engagement of students in minds-on activities which yield immediate feedback through discussion with peers and/or instructors. Such methods have been shown to consistently produce better conceptual understanding over traditional passive lectures in physics teaching (Hake, 1998, 2002). It is possible that the social interaction and active learning are responsible for the positive effect of ILDs, and not the cognitive conflict generated by the discrepant event.

Conclusions

1. We know that chemistry demonstrations are fun to do, provide concrete examples of abstract concepts, and are a potential source of anomalous data that can trigger conceptual change (Bodner, 2001). However, much of this potential is lost if a demonstration is simply carried out in front of a passive audience. To be effective, lecture demonstrations have to engage all students in activities such as prediction and discussion.
2. Cognitive conflict is not a magic wand that can solve all of the difficulties associated with conceptual change. It helps many students to achieve a high level of conceptual change, but it also fails for many others. Still, ILDs can provide a perfect setting for supporting classroom discussions, without emphasizing confrontation with students' prior conceptions (Ashkenazi and Weaver, 2007). Therefore, both supporters of cognitive conflict strategies and their opponents (such as Smith et al., 1993) can find use for them.
3. ILDs can be used without prediction sheets, just by collecting students' votes with an electronic classroom response system (for example: Wood and Brayfogle, 2006). However, we encourage instructors who use ILDs to ask students to write down their predictions, and collect the prediction sheets at the end of class. Not only does this make students more committed to their prediction, it is also an invaluable source of information for the instructor. It is a rare opportunity for lecturers in large classes to get to know their students' ways of thinking in a very detailed way. Such use of ILD combines its effectiveness as a teaching method for enhancing conceptual change, with its helpfulness as a research instrument for exploring this process and improving one's practice as a teacher.

Acknowledgement

We would like to thank Heather Sklenicka of Rochester Community and Technical College for her helpful comments.

References

Ashkenazi G. and Weaver G.C., (2007), Interactive lecture demonstrations as a context for classroom discussion : effective design and presentation, *Chemistry Education Research and Practice*, **8**. 186-196.

Bodner G.M., (2001), Why lecture demonstrations are 'exocharmic' for both students and their instructors, *University Chemical Education*, **5**, 31-35.

Chinn C.A. and Brewer W.F., (1998), An empirical text of taxonomy of responses to anomalous data in science, *Journal of Research in Science Teaching*, **35**, 623-654.

This journal is © The Royal Society of Chemistry

Crouch C.H., Fagen A.P., Callan J.P. and Mazur E., (2004), Classroom demonstrations: learning tool or entertainment? *American Journal of Physics*, **72**, 835-838.

Driver R., Guesne E. and Tiberghien A., (1985), Children's ideas and the learning of science, in R. Driver, E. Guesne and A. Tiberghien (eds.), *Children's ideas in science*, Open University Press, London, UK, 1-9.

Gobert J.D. and Buckley B.C., (2000), Introduction to model-based teaching and learning in science education, *International Journal of Science Education*, **22**, 891-894.

Galili I., Goldberg F. and Bendall, S., (1993), Effects of prior knowledge and instruction on understanding image formation, *Journal of Research in Science Teaching*, **30**, 271–303.

Galili I. and Hazan A., (2000), An Influence of historical oriented course on students' content knowledge in optics evaluated by means of facets – schemes analysis, *Physics Education Research American Journal of Physics Supplying*, **68**, 3-15.

Hake R.R., (1998), Interactive engagement versus traditional methods: A six thousand students survey of mechanics test data from introductory physics course, *American Journal of Physics*, **66,** 64-74.

Hake R.R., (2002), Lessons from the physics education reform effort, *Conservation Ecology*, **5** (2), article 28, http://www.consecol.org/vol5/iss2/art28..

Kang S., Scharmann L.C. and Noh T., (2004), Reexamining the role of cognitive conflict in science concept learning, *Research in Science Education*, **34,** 71-96.

Limon M., (2001), On cognitive conflict as an instructional strategy for conceptual change, *Learning and Instruction*, **11,** 357-380.

Mason L., (2001), Responses to anomalous data on controversial topics and theory change. *Learning and Instruction*, **11**, 453-483.

Nussbaum J., (1985), The particulate nature of matter in the gaseous phase, in R. Driver, E. Guesne and A. Tiberghien (eds.) *Children's ideas in science*, Open University Press, London, UK, 124-144.

Pfundt H. and Duit R., (2000), *Bibliography: students' alternative frameworks and science education*, University of Kiel Institute for Science Education, Kiel, Germany.

Posner G.J., Strike K.A., Hewson P.W. and Gertzog W.A., (1982), Accommodation of a scientific conception: toward a theory of conceptual change, *Science Education*, **66**, 211-227.

Shepardson D.P. and Moje E.B., (1999), The role of anomalous data in restructuring fourth graders' frameworks for understanding electric circuits, *International Journal of Science Education*, **21**, 77–94.

Sokoloff D.R. and Thornton R.K., (1997), Using interactive lecture demonstrations to create an active learning environment, *The Physics Teacher*, **35**, 340- 347.

Strike K.A. and Posner J.G., (1992), A revisionist theory of conceptual change. In R. Duschl, and R. Hamilyon (Eds.), *Psychology of science cognitive psychological and educational theory and practice*, Albany, NY pp.147-176.

Smith J.P., diSessa A.A. and Roschelle J., (1993), Misconceptions reconceived: a constructivist analysis of knowledge in transition, *The Journal of the Learning Sciences*, **3**, 115-163.

Tirosh D., Stavy, R. and Cohen S., (1998), Cognitive conflict and intuitive rules, *International Journal of Science Education*, **20**, 1257–1269.

Wood C. and Brayfogle B., (2006), Interactive demonstrations for mole ratios and limiting reagent, *Journal of Chemical Education*, **83**, 741-748.

This journal is © The Royal Society of Chemistry