

This is an electronic version of the article published in: Hulshof, C. D., Eysink, T. H. S., Loyens, S., & de Jong, T. (2005). ZAPs: Using interactive programs for learning psychology. *Interactive Learning Environments*, 13 (1&2), 39-53. © 2005 Taylor & Francis. *Interactive Learning Environments* is available online at: <http://www.informaworld.com/openurl?genre=article&issn=1049-4820&volume=13&issue=1&spage=39>

ZAPs: Using interactive programs for learning psychology

Casper D. Hulshof*, Tessa H. S. Eysink, Sofie Loyens^o & Ton de Jong

Faculty of behavioural sciences
Department of Instructional Technology
University of Twente
PO Box 217
7500 AE Enschede
The Netherlands

^o Faculty of Social Sciences
Department of Psychology
Erasmus University Rotterdam

* Corresponding author.

Tel.: +31-53-4893562/4894359; Fax: +31-53-4892895

E-mail: c.d.hulshof@utwente.nl

ZAPs: Using interactive programs for learning psychology

Abstract

ZAPs are short, self-contained computer programs that encourage students to experience psychological phenomena in a vivid, self-explanatory way, and that are meant to evoke enthusiasm about psychological topics. ZAPs were designed according to principles that originate from experiential and discovery learning theories. The interactive approach that is offered invites students to engage in subject matter through exploration, experience, and discovery of psychology. In an empirical study the effectiveness of different ZAPs for teaching psychology was examined. A group of students who worked with complete ZAPs was compared to a control group who worked with a ZAP from which the 'activity' component was removed. Posttest results showed that the control group outperformed the experiment group. However, on a retention test the differences between the groups disappeared. The results show that, on the long run, relatively good learning effects may be expected from working with ZAPs.

Introduction

Interactive software has become an established part of almost any standard curriculum. At the level of university education, this is especially true for the 'hard' sciences, notably physics and chemistry. In both disciplines, extensive use is made of computer simulated laboratories to enable students to explore phenomena that are difficult or impossible to reproduce in the real world.

There is a large body of literature that supports the claim that interactive learning fosters learning processes that are hard to obtain with more traditional teaching methods. Traditional teaching methods, which are also called expository, focus on acquiring factual knowledge. Examples of these are attending lectures and studying textbook material. A different view of learning, already expounded by Jerome Bruner in 1961 (Bruner, 1961) focuses on gaining understanding of a knowledge domain through mindful construction of 'personal' knowledge. Two main recent undercurrents in educational science can be distinguished that underlie the ideas set forth by Bruner. These are *experiential learning* and *discovery learning*.

Experiential learning refers to knowledge acquisition as a process of knowledge construction (Kolb, 1984). Learners themselves play a role in learning processes by actively engaging in the learning material. The theory of experiential learning stresses the need to start learning from concrete experiences that can be reflected upon. It is only after having had these experiences that a learner can make an abstract conceptualization of what it all meant. An experience involves a learner performing actions in a situation and observing the effect of these actions. Reflection means understanding the results of actions in a particular situation in such a way that predictions can be made regarding the outcome of similar actions in similar situations. Abstract conceptualization means gaining understanding of a general principle, so that predictions can be made about the effect of other actions in other situations. Although it may sound as a very abstract process, experiential learning may be one of the most concrete types of learning. In daily life we continuously experience new situations, from which we abstract information that

we apply in other situations. Experiential learning has relevancy to many different domains, the domain of psychology is no exception. In a way, the psychological processes that we deal with in everyday life can be described as experiential processes. In other situations, such as classroom situations, computers can be useful tools in eliciting experiential learning processes. Interactive software is instrumental in achieving experiential learning on the computer. This may be even more true for the psychological domain than for physics and chemistry. Psychology has human beings as its subject, that is, it deals with ourselves. This means that psychological phenomena may be best understood by subjectively experiencing them and by reflecting upon that experience. One way the computer can foster experiential learning of psychology is by presenting phenomena and having learners interact with them. Examples are visual illusions such as the well-known Ponzo illusion (shown in Figure 1), of which different boundary conditions can be explored. In the description of ZAPs below, this point will be more thoroughly explicated. Another way to use the computer for experiential learning is by using a computer simulation. Experiential learning in the context of a computer simulation is also known as ‘discovery learning’ (De Jong & Van Joolingen, 1998).

 Insert Figure 1 about here
 #####

Discovery learning requires students to devise one or more hypotheses and test their validity by performing a number of relevant experiments. Over the years, the computer has proven itself invaluable in the creation and testing of various types of simulations to foster discovery learning about a conceptual domain (de Jong & van Joolingen, 1998). The focus on conceptual domains sets restrictions on the type of domains that can be simulated. A domain can only be simulated if there is actually something to simulate, that is, a *model* of the domain. For that reason, most educational simulations have been restricted to physical, chemical, or economical domains. As stated above, the method of discovery learning can also be relevant to learning about psychological phenomena. Many different phenomena, ranging from biological to cognitive processes, can be described in terms of a model. An example is the gate-control theory of pain sensation (Melzack & Wall, 1965). Common descriptions of this theory involve different types of fibers and synaptic ‘gates’ that together shape different types of pain sensation. The gate-control theory describes the relation between phenomena that occur at the physical level of neural interaction and conscious coping strategies at the psychological level. Such a description takes a lot of effort to understand. Also, a verbal description necessarily misses some optional possibilities or issues that one may be interested in. A functional model of the theory helps overcome these problems. Experimenting on a model of the gate-control theory supports deep understanding of a complex phenomenon. In addition, such models may be used to show how a computational description reveals the relation between biological and psychological levels of description.

In sum, it can be argued that it is possible to apply both the theory of experiential learning and discovery learning to the domain of psychology. This involves creating learning material that offers a direct experience with a phenomenon, or the possibility to perform experiments to discover an underlying model. In the ZAP project, the goal was

to apply both principles to the design of interactive software modules (called ‘ZAPs’) in the context of the study of psychology. ZAPs are targeted for use by first-year psychology students. This means that for those students, ZAPs are designed to be easy to use, simple to understand, yet rich enough in content to contain a full ‘experience’. The design and structure of ZAPs is only discussed in brief here. A more thorough description, in which the relation between implications from experiential and discovery learning theories and the construction of ZAPs is more fully explored, is discussed in a forthcoming article (Hulshof, Eysink, & De Jong, 2004).

Figure 2 shows an example screen of a ZAP. In this case, the ZAP is about the topic of classical conditioning, a basic type of learning that was first systematically investigated in the early twentieth century by Ivan Pavlov.

 Insert Figure 2 about here
 #####

The left part of Figure 2 shows a number of different components (‘Introduction’, ‘Discovery’, ‘Theory’, and ‘Further Info’). These four components are part of all ZAPs. Three of the four components contain text and pictures. The goal of the text parts is to offer a short introduction to a topic to students (in the Introduction component), to offer a theoretical explanation (in the Theory component), and to offer further explanations and comment on related phenomena (in the Further Info component). The heart of all ZAPs is formed by the ‘Activity component’ (called ‘Discovery’ in Figure 2). ZAPs can be divided into three different types, based on the type of activity they contain: ‘experience ZAPs’, ‘discovery ZAPs’, and ‘experiment ZAPs’. In experience ZAPs, learners directly experience in what way they themselves react to certain psychological phenomena. Examples of phenomena that can be experienced are illusions as displayed in Figure 1, but also higher mental processes such as biases in reasoning or problem solving methods. In discovery ZAPs, learners take the role of experimenter in a virtual laboratory. By setting up and performing experiments, a theory of the model underlying behavior of a virtual subject can be derived. Example phenomena that can be modeled in a discovery task are the already mentioned gate-control theory of pain and classical conditioning, but also human behavior such as the behavior of people whose corpus callosum has been cut (also called ‘split brain patients’). In experiment ZAPs, learners do not perform experiments on someone else, instead they are participants themselves in a (classic) psychological experiment. An experiment consists of several trials in which the participant’s task is to respond in various ways to stimuli that are presented. The end result is a data set that shows by comparison of different experiment conditions the way a theory of cognitive processing can explain results obtained from an experiment. Examples are the mental rotation experiments by Shepard and Metzler (1971) which illustrate the relation between perception and mental imagery, and the Stroop effect (Stroop, 1935) which illustrates the interference between controlled and automated tasks.

During the development of ZAPs, prototype versions were extensively evaluated. Based on this evaluation, improvements were made to the textual components, and revisions were made to the activity component. The evaluations led to many favorable remarks, and the ZAPs have so far been very well received by both students and teachers.

However, an important question remains: do ZAPs contribute positively to learners' knowledge about psychological topics? To answer this question, an experiment was conducted. The current experiment focused on the contribution of the activity component of ZAPs (experience, discovery, or experiment) to the acquirement of psychological knowledge. Since the activity component takes up a central position in a ZAP, it was expected that learners would benefit from interacting with it. A number of studies in which the effectiveness of instruction based on experiential and discovery learning was studied have shown that its result is insight and deep understanding of a topic, not necessarily more knowledge than would be obtained in a more traditional setting (e.g., Swaak & De Jong, 1996). The implication is that the gain from working with the activity component of a ZAP should be measured by looking at insight into and understanding of the psychological phenomenon that is covered by the ZAP. A ZAP without the activity component only consists of a structured text on a psychological topic. It was expected that studying the text would lead to a gain of factual knowledge about a psychological topic, as opposed to real understanding. In sum, positive learning outcomes from working with ZAPs were expected. It was expected that factual knowledge would be the same and insight in the psychological phenomena would be higher for learning with complete ZAPs compared to learning with ZAPs from which the activity component was removed. Furthermore, it was expected that, on the long term, students who were presented with complete ZAPs would better maintain their acquired knowledge and insight than students who received incomplete ZAPs. These results were expected for all three ZAP types (experience, discovery, and experiment).

Method

Participants

In total, 71 participants took part in the experiment: 30 participants were students enrolled in higher vocational education, and 41 participants were university students enrolled at the University of Rotterdam, both located in The Netherlands. Students had not yet received any formal instruction on the topics covered in the different ZAPs. For the higher education students, participation in the experiment was scheduled as part of the school's first introduction week. Because the university students volunteered for the experiment, they were financially compensated for their participation.

Design

A randomized two-group pre-post-retention test design was used for the experiment. A pretest was included, but it was differently constructed from the posttest and retention test so it was not possible to compare results on the pretest to the other tests. The reason for using a different pretest was to prevent participants from being influenced in the experiment by the questions. Two conditions were used, an experimental condition in which participants worked with complete ZAPs during the experiment (that is, texts and the interactive component), and a control condition in which participants only read the text components without the interactive component.

Participants were randomly assigned to the experimental condition ($n=35$) or the control condition ($n=36$). Participants in the experimental condition worked with complete ZAPs (that is, ZAPs consisting of a number of text components and an activity

component). Participants in the control condition worked with ZAPs consisting of text components without the activity component. Participants were neither made aware of the fact that they were assigned to a particular condition, nor of the difference between the conditions. All participants worked with six different ZAPs during the experiment. To control for sequence effects or fatigue, the sequence in which participants worked with ZAPs differed for each person. It was not possible for participants to work through the ZAPs in any other sequence than the one that was offered to them.

Materials

From the ZAPs that were produced as part of the project, six were selected for use in the present experiment. The selected ZAPs were representative for the complete collection. For each type of ZAP (experience, discovery, or experiment) two representative ZAPs were selected. Table 1 lists the titles of the ZAPs that were selected for each type. A short description of the contents of the ZAPs is also given in the table.

 Insert Table 1 about here
 #####

The largest difference between ZAPs in the experiment condition and the control condition was the presence of an activity component in the former. In addition, to cope with the absence of an activity from the control condition minor textual modifications were made to the text to remove any references to the activity. Also, instead of actually working with an activity, participants in the control condition read a short explanation that described the original experiment on which the ZAP was based. In both conditions, the ZAPs were opened within a web browser. There were no technical difficulties in opening or running the ZAPs during the experiment.

Tests

In total, three tests were administered to participants: a pretest, a posttest, and a retention test. All three tests covered the six topics dealt with in the ZAPs plus an additional topic about an additional seventh ZAP which was not used in the experiment. The seventh topic (on the ‘serial position’ memory effect) was used in order to control for experimental artifacts.

The pretest consisted of 14 open-answer questions, 2 for each topic, the aim of which was to measure the extent to which participants had prior knowledge about the topics that were covered in the different ZAPs. An example item from the pretest is ‘What is the *corpus callosum*, and what is its function?’. It was expected that participants would have only little prior knowledge about psychology, since they had just entered their first year of higher education. Therefore, it was specifically stated in the instructions to the test that it would not matter if they answered ‘no idea’ to the questions. This was done to prevent participants from becoming frustrated or unmotivated. The test items were constructed in such a way that the probability of focusing participants’ attention to specific parts of the ZAPs would be minimal. As a consequence, the style of the pretest was different from the posttest. The posttest consisted of 35 open questions. There were 5 questions for each of the 6 ZAPs covered in the experiment, and 5 additional questions about the additional

seventh topic. The questions of the posttest concerned the topic of each ZAP used in the experiment. The 5 questions for each topic were divided over 2 ‘knowledge questions’ about specific factual content, and 3 ‘insight questions’ that involved predicting imaginary variations on the situation that was covered by the activity component of the ZAP. Knowledge questions covered the textual components of a ZAP. An example knowledge question (about the 2-4-6 Task) is: ‘What is meant by *inductive reasoning*? Please provide an example of inductive reasoning’. Most insight questions were longer than the knowledge questions. They usually described a situation that might occur in the activity component, and asked participants to imagine and predict what would happen if the situation was modified in some way. An example insight question about classical conditioning is shown in Figure 3. All insight questions were constructed in such a way that they could be answered by participants in both the experiment and the control condition.

 Insert Figure 3 about here
 #####

Answers to the open questions were scored afterwards. The 5 questions for each ZAP added up to 10 available points. The division for each ZAP was 4 points for the knowledge questions, and 6 points for the insight questions. Division of points over the questions varied with the content.

To measure the stability of knowledge gained in the experiment, a retention test was administered approximately ten days after the experiment. The retention test was similar to the posttest. About half the questions from the knowledge test had been slightly modified, to avoid complete equivalence between the tests. Again, participants answered 35 questions, part knowledge questions and part insight questions.

Process measures

While navigating through the ZAPs, all operations participants carried out (as measured in ‘mouse clicks’) were unobtrusively registered by the computer. This made analysis of the time subjects spent on different components of a ZAP possible. Because the navigation log data was combined with student identification data, it was also possible to compare the experiment and control condition with respect to navigation through ZAPs.

Procedure

The experiment was divided over two sessions. In the first session, participants worked individually at the computer. The session was completely web-based, which means that both the tests and the ZAPs were opened within a web browser. All participants first received an individual login code, and used this code to start the experiment session. After filling out a form with some personal information (name, age, prior education), the knowledge pretest started. Because most participants had little prior knowledge about psychological topics, completing the pretest took only a short time. Immediately after the pretest, participants started working with the first of 6 ZAPs. Depending on the login code, each participant worked with the ZAPs in a different

sequence. To prevent participants from skipping a ZAP, they had to spend a minimum time of 3 minutes on working with a ZAP before they were allowed to continue to the next one. Early exits from a ZAP were not possible: attempts to do so were met with a warning, which asked the student to spend at least a minimum time with the ZAP. After completing all 6 ZAPs, the posttest was administered to participants. The questions on the posttest were delivered in the same order as the sequence of ZAPs that a participant had worked with. For example, a participant who had first worked with the cognitive dissonance ZAP and had finished with the split-brain ZAP would first receive 5 questions about cognitive dissonance and end with 5 questions about split-brain. The session ended as soon as the posttest was completed. The complete session took about two hours.

The second session took place approximately 10 days after the first. In this session, participants did not work with the computer. Instead, a retention test was administered as a pencil-and-paper test. The sequence in which the questions were printed on the question sheet varied: for each participant it was the same as the sequence of questions on the posttest.

Results

Two types of results are of interest in the context of the present experiment: students' performance on the tests (pretest, posttest, and retention test) and process data. Analysis showed no difference between the two student populations that participated in the experiment (from higher education and university education), so only results for the complete group are described in this section. Test results are described first, followed by process results.

Pretest. As was expected beforehand, students experienced much trouble in answering the pretest questions. Most questions were answered by almost all students with 'Don't know', or 'No idea'. Because of this it was not possible to compute a reliable value of Cronbach's α . The only question that showed relatively good answers were the ones on the topic of classical conditioning. The reason for this may be that this topic is regularly taught in Dutch secondary schools as part of the biology curriculum. Other results from the pretest revealed a number of misconceptions about psychological phenomena. For example, cognitive dissonance was seen as a pathological phenomenon (which is not the case), and split-brain patients were regarded as suffering from a multiple personality syndrome (which they are not). In general, the results show that prior knowledge about the different subjects that were subsequently covered in the ZAPs was very poor. Furthermore, the results on the pretest show that the experiment and the control condition are comparable, as they did not differ ($F_{1,69} = .05$, n.s.).

Posttest and retention test. Cronbach's α for the complete posttest was .81. Cronbach's α for the complete retention test was .73. The correlation between performance on the posttest and the retention test was .72. To measure differential learning gains from working with the ZAPs, posttest and retention test results for the experiment and control condition were compared. In Table 2 the results are shown, divided over the 3 ZAP types that were used (for each type, 2 ZAPs were used in the experiment, as shown in Table 1). This resulted in an average test score that could range from 0 to 10. On the posttest, the control condition and experiment condition scored 4.81 (SD=1.45) and 3.79 (SD=1.36), respectively. On the retention test, the averages are 4.16

(SD=1.25) for the control condition, and 3.87 (SD=1.09) for the experiment condition. A statistically significant interaction effect between the test (posttest or retention test) and condition (experiment or control) was found ($F_{2,57} = 4.40, p < .05$).

 Insert Table 2 about here
 #####

The results show that participants' total score in the experiment condition was lower on the posttest than that of participants in the control condition ($F_{1,67} = 10.18, p < .01$). This difference was caused by both different performances on the knowledge questions ($F_{1,67} = 9.13, p < .01$) and on the insight questions ($F_{1,67} = 6.97, p < .01$). Differences between the conditions on the posttest are found for all three ZAP types, experiment ZAPs ($F_{1,69} = 5.36, p < .05$), discovery ZAPs ($F_{1,68} = 7.54, p < .01$) and experience ZAPs ($F_{1,67} = 5.59, p < .05$). Results on the retention test show different results. No significant difference appeared between the experiment and control condition ($F_{1,61} = 1.28, n.s.$). Also, no differences are found for the knowledge and insight questions, nor for the different ZAP types. To make the test results more clear, Figure 4 shows a graph of the comparison between the two conditions on the posttest and the retention test. The test scores that are displayed in the figure are average test scores per ZAP (that is, knowledge and insight questions, averaged for all ZAPs).

 Insert Figure 4 about here
 #####

As can be seen from Figure 4, the control condition shows a significant decrease in performance from the posttest to the retention test ($F_{1,32} = 10.83, p < .01$), whereas the results of the experiment condition remain the same from posttest to retention test ($F_{1,25} = .26, n.s.$).

Posttest and retention test items of the seventh topic. As described in the Method section, the posttest and the retention test also consisted of items on a topic on which students did not receive a ZAP. This topic was used in order to control for experimental artifacts. The results on the questions about this topic were, as was expected, low. There was no difference between the experiment condition and the control condition ($F_{1,59} = .01, n.s.$), and the scores on the posttest and the retention test remained the same for both the control condition ($F_{1,32} = .00, n.s.$) and the experiment condition ($F_{1,27} = .01, n.s.$). When the results on this topic are divided between knowledge questions and insight questions, a difference between scores on these types of questions becomes clear, though. On the knowledge questions students score low and their score on this topic is much lower than their score on the topics they received ZAPs on. However, on the insight questions the difference between scores on topics they received ZAPs on and scores on the seventh topic is very small.

Process results. For all participants, time spent on reading a text component or on working with the activity component (in the experiment condition) was registered. In ZAPs of the experiment type, time was also registered for the extra data component that

this type contained. When a participant had viewed a component for less than 5 seconds, the result was removed from the analysis. The results are displayed in Table 3.

 Insert Table 3 about here
 #####

As can be seen from the table, students in the experiment condition on average spent 8:45 minutes working with a ZAP, compared to 5:42 minutes for students in the control condition. The time difference can be explained by the presence of an activity component in the experimental condition, which was absent in the control condition. However, it can be observed in the table that there are consistent differences between the conditions in the average time spent reading the textual components. In the control condition, participants spent more time reading the texts than students in the experiment condition. When the average time spent on ZAPs is divided over the three different types (experience, discovery, or experiment), it can be seen that there are large differences in the time participants spent working with each. For experience ZAPs and discovery ZAPs differences are the largest: participants in the experiment condition spent much longer working with experience ZAPs than with discovery ZAPs. For the experiment ZAPs the data is less clear because the experiment in the activity component takes a minimum time to complete for all participants. The short average time spent on the activity component of the discovery ZAPs indicates that participants may have had trouble understanding the rich content of the activity in the discovery ZAPs. Process data confirmed this hypothesis, because it showed that most participants performed only a few experiments in these ZAPs.

Discussion

Our expectations were to find no differences between conditions for the knowledge questions of the posttest and better scores on the insight questions of the posttest for the experiment condition. Furthermore, better conservation was expected on the long run for the experiment condition.

To start with the latter, the results partly confirmed the hypothesis. The interactive component in which students actively engaged with the subject matter was meant to lead to good understanding and integration of knowledge, leading to prolonged conservation. The results showed that the scores of students in the control condition decreased from posttest to retention test, whereas scores of the experiment condition remained the same over time. In the interval between the experiment and the retention test, students in the experimental condition better maintained their knowledge than students in the control condition. The combination of the textual parts with the activity in the experiment condition may support better understanding of psychological information, which means that memory of this information degrades more slowly over time.

The expectations concerning the posttest administered immediately after the experiment were not confirmed by the results. The interactive component was especially meant to improve insight in the topic of the ZAP, but students in the experiment condition performed less on the insight questions than students in the control condition.

However, the scores on the insight questions of the seventh topic show that the insight questions that were used might not have measured insight as was intended. The idea behind the design of the insight questions was to measure deep understanding of the subject matter. This was operationalised by asking students to imagine a variation of the standard situation that was used in the ZAP activity, and to answer a question based on this hypothetical situation. It was thought that, in order to come up with an answer, students needed deep understanding of the subject matter. However, students' scores on the insight questions of the seventh topic, the topic on which they had not received any instruction, were almost as high as the scores of students on the insight questions of topics on which they had received instruction. It can be hypothesized that answering the insight questions did not necessarily require deep understanding of the subject matter. The answers could possibly also be generated by correct reasoning. Most insight questions involved imagining the situation, after which it may have been possible to have reason follow its logical course. Possibly, insight was not necessary to successfully answer the insight questions, nor was instruction into the topic.

The finding that participants in the experiment condition scored lower on knowledge questions of the posttest than participants in the control condition was also not in accordance with prior expectation. The knowledge questions concerned knowledge which was presented to the students in the textual components of the ZAP. As these textual components were the same for both conditions, no differences were expected on the posttest between the experiment and the control condition. However, the fact that the experiment condition scored lower may, at least in part, be explained by the setup of the study. In the experiment condition, the students worked with a number of normal, complete ZAPs. Performance of this group was compared to a control condition, in which students were assigned only one task: to carefully read and understand a number of texts. Students in the first group faced the dual task of interpreting texts and carrying out an interactive assignment. It may well be that students in this condition saw the purpose of the text parts primarily as an illustration to the activity component. That would mean that these students focused less on the factual content of the texts, and more on the information pertaining to the task at hand. Evidence for this explanation can be observed by analysis of the time spent on each ZAP component: time spent on reading textual components—that were equivalent for both conditions—was systematically shorter for the experiment condition than for the control condition. On the whole, students in the control condition were given a more simple task than students in the experiment condition, since their only task was to read texts. They were aware of the fact that afterwards, they would receive questions on the content of these texts. In retrospect, the difference in tasks assigned to students in both conditions may be responsible for causing a better performance on the knowledge questions of the posttest for the control condition compared to the experiment condition.

Conclusion

ZAPs are positively welcomed by both students and teachers. Psychological phenomena that students usually only read about can be fully experienced and experimented with. In this respect, they succeed in evoking enthusiasm about a variety of psychological topics. Although the present study did not find beneficial learning gains for

students who used complete ZAPs immediately after instruction, compared to results for students who used ZAPs without the interactive component, results on the long run were positive in the sense that students who used the complete ZAPs better conserved their knowledge. It may be hypothesized that this difference in conservation of knowledge will also be found on a long-term retention test. It should be kept in mind that ZAPs are originally meant to be integrated in existing courses. The small modules are highly suitable for this purpose. One or more ZAPs can easily be used in addition to existing learning materials. The value of ZAPs lies in the interaction they add to otherwise passively consumed texts, the combination of factual information with an attractive and interactive example, the possibility of easily integrating ZAPs in existing courses, and the added variety that makes a ZAP worthy of repeated visits. All these opportunities that ZAPs offer were not taken advantage of in the current study. Future studies should focus more on these aspects, so that the added value of ZAPs in existing curricula comes more to its right.

Acknowledgements

The development of ZAPs took place in the context of the ZAP project, which was supported by a grant from the Stichting SURF. The ZAP project team consisted of Henk Schmidt, Wilco te Winkel, Sofie Loyens, Eveline Osseweijer and Manon de Jong of the Erasmus University Rotterdam, and Jules Pieters, Jan Oosterhuis, Jakob Sikken, Marita Wesselink, Ben Reimerink, Ton de Jong, Casper Hulshof and Tessa Eysink of the University of Twente.

References

- Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, 31, 21-32.
- Hulshof, C. D., Eysink, T. H. S., & De Jong, T. (2004). *The ZAP Project: Designing Interactive Computer Tools for Instruction in Psychology*. Unpublished Manuscript.
- De Jong, T., & Van Joolingen, W. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179-202.
- Kolb, D. A. (1984). *Experiential Learning*. Englewood Cliffs, NJ: Prentice-Hall.
- Melzack, R., & Wall, P. D. (1965). Pain mechanisms: A new theory. *Science*, 150 (3699), 971-979.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703.
- Swaak, J., & De Jong, T. (1996). Measuring intuitive knowledge in science: The development of the What-if test. *Studies in Educational Evaluation*, 22, 341-362.
- Stroop, J. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, 18, 643-662.

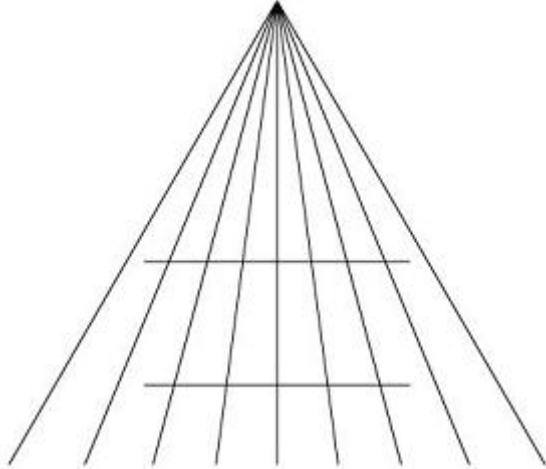



Figure 1. Example of the Ponzo illusion. Which horizontal line is the longest?

ZAP **Classical conditioning** Print version

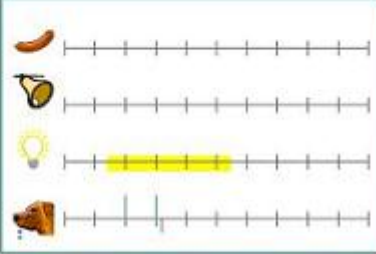
Duration: 15 minutes

Introduction
Discovery
 Theory
 Further info

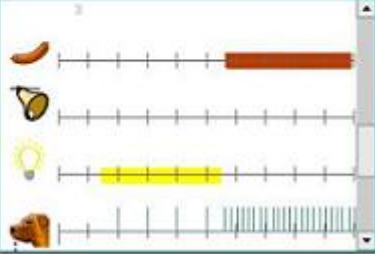
Configuration



Experiment



Results of previous experiments

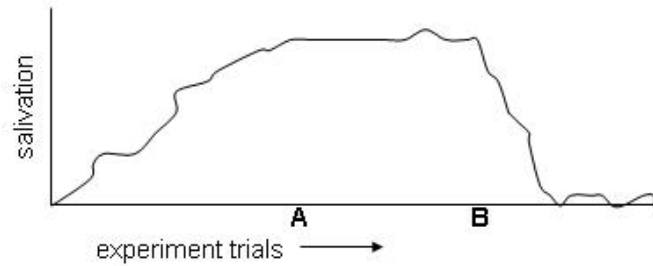


help restart further

©ZAP 2004

Figure 2. Example of the activity component of a ZAP on classical conditioning

You are performing learning experiments on a dog. Let's call every time you present something to the dog and measure how much the dog salivates a 'trial'. The graph below shows results from a large set of trials. From the beginning to time A, each trial the dog was presented with meat and the sound of a bell. From time A on, the dog was presented with both meat and a bell sound, and in addition a light was flashed. The procedure of presenting three stimuli was repeated a number of times, until time B. From that moment on only the lamp was flashed during trials. The graph shows the amount of salivation for each time period.



The question is: why is there a lack of salivation after time point B?

Figure 3. Example insight question (on classical conditioning)

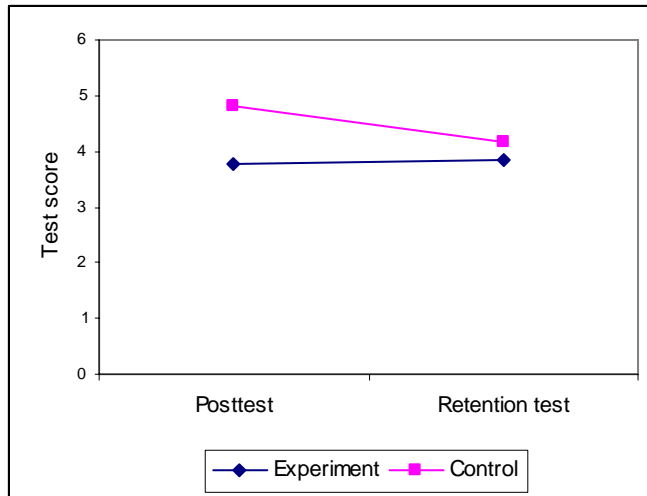


Figure 4. Comparison between posttest and retention test performance for both conditions, averaged per ZAP. Test scores could range from 0 to 10

Table 1. Description of ZAPs used in the experiment

<i>Title</i>	<i>Type</i>	<i>Description</i>
2-4-6 Task	} Experience	The effect of the confirmation bias in inductive reasoning
Cognitive dissonance		Unpleasant situations and ways of resolving them
Classical conditioning	} Discovery	Different classical basic learning phenomena
Split-brain patient		Different behavior of the left and right part of the brain
Spatial cueing	} Experiment	Inducing implicit expectations by cueing
Stroop effect		Interference effects in judging colors

Table 2. Percentage of correct answers on posttest and retention test, both conditions, split by ZAP type

ZAP Type	Test	Question type	Condition	
			Experiment	Control
Discovery	Posttest	Knowledge	40,5%	52,8%
		Insight	36,8%	46,8%
	Retention test	Knowledge	25,0%	34,8%
		Insight	35,7%	44,7%
Experience	Posttest	Knowledge	31,3%	47,5%
		Insight	38,7%	47,5%
	Retention test	Knowledge	41,0%	39,5%
		Insight	41,3%	46,2%
Experiment	Posttest	Knowledge	19,5%	26,0%
		Insight	53,0%	61,8%
	Retention test	Knowledge	12,0%	12,5%
		Insight	62,8%	59,3%
Overall	Posttest	Knowledge	30,4%	42,1%
		Insight	42,8%	52,0%
	Retention test	Knowledge	26,0%	28,9%
		Insight	46,6%	50,1%

Table 3. Average time (minutes and seconds) spent on ZAP component, both conditions

Condition	Component	All ZAPs	Experience	ZAP Type	
				Discovery	Experiment
Experiment condition (complete ZAP)	Introduction	1:07	1:10	1:15	0:57
	Instruction	0:52	1:01	0:43	0:52
	Activity	4:14	6:55	2:31	3:17
	Data	0:45	--	--	0:45
	Theory	1:28	1:19	1:53	1:11
	Further info	0:59	1:25	0:42	0:51
	Total	8:45	11:50	7:03	7:54
Control condition (text only)	Introduction	1:25	1:30	1:34	1:12
	Instruction	1:04	1:38	0:35	0:58
	Theory	1:58	1:54	2:21	1:39
	Further info	1:01	1:51	0:56	1:01
	Total	5:42	6:52	5:25	4:51