Abstract: Discovery-based learning designs incorporating active exploration are common within instructional software. However, researchers have highlighted empirical evidence showing that ‘pure’ discovery learning is of limited value and strategies which reduce complexity and provide guidance to learners are important if potential learning benefits are to be achieved. One approach to reducing complexity in discovery learning is limiting the range of possible actions for the learner to ensure that they do not undertake exploratory activities leading to confusion. This article reports on a study in which the learning outcomes from two learning conditions using computer-based simulations were compared. One condition allowed exploration through manipulation of simulation parameters, while the other allowed observation of simulation output from preset parameters, the latter condition designed to limit the complexity of the task. Learning outcomes for the 158 university student participants were assessed via pre-tests and post-tests of conceptual understanding. Students’ exploration activities were recorded and their strategies subsequently coded as either systematic or unsystematic. The results showed that when compared with observation, systematic exploration resulted in learning benefits, while unsystematic exploration did not. These results have implications for the design of discovery learning tasks and instructional guidance within computer-based simulations.

URLs:

FT: http://dx.doi.org/10.1080/09523987.2014.977009

The impact of students’ exploration strategies on discovery learning using computer-based simulations

Associate Professor Barney Dalgarno (corresponding author)

School of Education, Charles Sturt University, Wagga Wagga, Australia

Locked Bag 588, Boorooma Street, Wagga Wagga NSW 2678 AUSTRALIA

bdalgarno@csu.edu.au
+61 2 6933 2305

Professor Gregor Kennedy

Centre for the Study of Higher Education, University of Melbourne, Australia

Level 1, 715 Swanston St, Parkville VIC 3010 AUSTRALIA

gek@unimelb.edu.au
+61 3 8344 8937

Associate Professor Sue Bennett

School of Education, University of Wollongong, Australia

The School of Education, Building 23.G21, University of Wollongong, Northfields Avenue, Wollongong NSW 2522 AUSTRALIA

sbennett@uow.edu.au
+61 2 4221 5738

Acknowledgements

Support for this project has been provided through a Charles Sturt University Small Grant and a University of Wollongong Centre for Research in Interactive Learning Environments Seed Grant. The authors would also like to acknowledge the contributions of Dr Terry Judd and Dr Michael Lew of the University of Melbourne to the development of the computer-based simulations.
The impact of students’ exploration strategies on discovery learning using computer-based simulations

Discovery-based learning designs incorporating active exploration are common within instructional software. However, researchers have highlighted empirical evidence showing that ‘pure’ discovery learning is of limited value and strategies which reduce complexity and provide guidance to learners are important if potential learning benefits are to be achieved. One approach to reducing complexity in discovery learning is limiting the range of possible actions for the learner to ensure that they do not undertake exploratory activities leading to confusion. This article reports on a study in which the learning outcomes from two learning conditions using computer-based simulations were compared. One condition allowed exploration through manipulation of simulation parameters, while the other allowed observation of simulation output from preset parameters, the latter condition designed to limit the complexity of the task. Learning outcomes for the 158 university student participants were assessed via pretests and posttests of conceptual understanding. Students’ exploration activities were recorded and their strategies subsequently coded as either systematic or unsystematic. The results showed that when compared with observation, systematic exploration resulted in learning benefits, while unsystematic exploration did not. These results have implications for the design of discovery learning tasks and instructional guidance within computer-based simulations.

Keywords: discovery learning, inquiry learning, computer-based simulations, instructional guidance

Introduction

The notion of discovery learning has its origins in the 1960s, with Jerome Bruner one of the first to articulate in detail the potential benefits of instructional approaches with discovery learning at their core (Bruner, 1961). There are a range of related learning design approaches which are similar to or draw on elements of discovery learning, including exploratory learning (Reilly, 1974; De Freitas & Oliver, 2006), inquiry learning (Kuhn, Black, Keselman & Kaplan, 2000; Rutherford, 1964), and Problem Based Learning (Barrows & Tamblyn, 1980). Discovery learning and related learning design approaches have theoretical support in cognitive interpretations of constructivism which can be traced to Piaget’s focus on active
knowledge construction through exploration and interaction with one’s environment (Duffy & Cunningham, 1996; Klahr & Nigam, 2004). The idea that learning involves active knowledge construction has particularly been used in support of inquiry-based learning approaches in the sciences, including discovery learning involving the use of computer-based simulations (De Jong & Van Joolingen, 1998).

In contrast to the well accepted theoretical support for discovery learning, a number of highly cited articles have questioned the value of discovery learning and particularly ‘pure’ discovery learning where little or no guidance is provided during the learning task. Mayer (2004), for example, reviewed three decades of research on discovery learning and concluded that in each case guided discovery learning was more effective than pure discovery learning. Kirschner, Sweller and Clark (2006), reached similar conclusions based on an argument grounded in current knowledge about cognitive architecture, expert-novice differences and cognitive load. Most recently, Alferi, Brooks, Aldrich and Tenenbaum (2011) in a meta-analysis of 164 empirical studies of learning using discovery-based approaches concluded that explicit instruction is more effective than unassisted discovery, while discovery enhanced by the inclusion of guidance during learning was more effective than both unassisted discovery and explicit forms of instruction.

Clearly there is a major disjunction between the focus on discovery within well accepted constructivist informed teaching approaches, and the available empirical results. Given this contention, there is a need for further studies which explore student learning processes and outcomes through discovery learning. This article reports on a study investigating how learners’ exploration strategies impacted on their learning outcomes while undertaking discovery learning activities using computer-based simulations.

**Background**
**Discovery learning**

Bruner (1961) in a frequently cited article on the benefits of discovery learning, distinguishes between instructional approaches which involve discovery or “finding-out-for-oneself” (pp. 22-23) and traditional expository approaches in which the key instructional decisions are made by the teacher. He argues in favour of discovery approaches on the basis that such approaches are more intrinsically motivating, lead to greater retention and recall, and help develop important inquiry and problem-solving skills. In an early review of research on discovery learning, Kersh and Wittrock (1962) state that “the term ‘discovery’ frequently describes a learner’s goal directed behaviour when he is forced to complete a learning task without help from the teacher” (p. 461). Alferi et al. (2011) highlight the absence of explicit instruction as a defining feature of discovery learning and add that “the target information must be discovered by the learner within the confines of the task and its material” (p. 2). In contemporary applications of discovery learning students typically explore printed or online learning resources (Steuter & Doyle, 2010), physical spaces or artefacts (Klahr & Nigam, 2004), virtual environments (Lee & Dalgarno, 2011), computer-based simulations (De Jong & Van Joolingen, 1998) or computer games (De Freitas & Oliver, 2006), with minimal guidance prior to or during the task, and are expected to discover key factual information or develop a conceptual understanding of key ideas as part of their exploration.

A key element of constructivist theories of learning, and one which underpins discovery learning and related instructional approaches, is the idea that each person forms their own knowledge representation, building on their individual experiences – an idea generally attributed to Piaget (1973). According to constructivist theory, this knowledge representation is constantly reviewed and revised as inconsistencies between the learner’s current knowledge representation and experience are encountered through active exploration (Bruner, 1962; Von Glasersfeld, 1984). Piaget (1973) explains the learning process in terms
of equilibration. Equilibration begins with the construction by the individual of their own internal knowledge representation, or in Piaget’s terms they accommodate their knowledge representation or schema to fit with their experience. Subsequent experiences that are consistent with this knowledge representation are then assimilated into this schema. New experiences that do not fit with their current knowledge representation result in a further accommodation of their schema to fit with this new experience. Clearly, such an account of the learning process, with its emphasis on constructing and reconstructing an individual knowledge representation through active exploration has a natural fit with the idea of discovery learning.

Piaget’s account of the learning process is generally well accepted although some researchers describe the process using different terminology (see Duffy & Cunningham, 1996). Others, especially those articulating a social rather than cognitive constructivist position, drawing particularly on Vygotsky (1978), argue that such an account is incomplete because it does not include social factors and particularly the role played by peers and teachers in supporting the learning and knowledge construction process (Phillips, 1995). Some critics of discovery learning have argued that unguided discovery is not a logical consequence of constructivist learning theories rather than arguing against Piagetian accounts of learning. Kirschner, Sweller and Clark (2006), for example, state that “the constructivist description of learning is accurate, but the instructional consequences suggested by constructivists do not necessarily follow” (p. 78). They argue instead for the use of direct instructional guidance as the most effective way to support learners in the knowledge construction process given what is known about limits in working memory capacities. A key implicit element of this argument is a distinction between behavioural activity and cognitive activity and specifically the idea that, firstly, not all behavioural activity leads to cognitive knowledge construction and that, secondly, cognitive knowledge construction can occur
without behavioural activity (see Alferi et al., 2011). Piaget himself warns that a common misuse of his ideas is “that which leads people to think that any ‘activity’ on the part of the student or child is a matter of physical actions, something that is true at the elementary levels but is no longer so at later stage, when a student may be totally ‘active’, in the sense of making a personal rediscovery of the truths to be acquired, even though this activity is being directed towards interior and abstract reflection” (Piaget, 1977; p. 714). This distinction between behavioural and cognitive activity is particularly pertinent to this study and the implications for the design of discovery learning activities using computer-based simulations are discussed further below.

**Computer-based simulations**

A popular manifestation of discovery learning approaches have been computer-based instructional simulations focussing on the learning of conceptual material in the sciences (De Jong & Van Joolingen, 1998). These simulations began to emerge in the 1990s, with resources like *Investigating Lake Iluka* (Harper, Hedberg & Brown, 1995) and the *Jasper Woodbury* series (Cognition and Technology Group At Vanderbilt, 1992), which allowed non-sequential inquiry-based exploration of a computer-based learning environment underpinned by the provision of holistic problem scenarios. At the time these resources emerged they were noteworthy because of their contrast to prevailing computer-based learning resources which largely consisted of programmed instruction or tutorials in which the learner would work through a linear sequence of instructional screens or drill-and-practice activities which provided an opportunity for learners to perfect their responses with immediate feedback.

Computer-based instructional simulations commonly employ ‘predict-observe-explain’ learning designs (see, for example, Monaghan & Clement, 1999). These designs require the learner to first “predict the outcome of some event” and “justify their prediction”,
then “describe what they see happen”, and finally “reconcile any conflict between prediction and observation” (White & Gunstone, 1992, p. 44). Similar learning designs have also been used under other labels, for example scientific discovery learning (see De Jong and van Joolingen, 1998) and learning by hypothesis-testing (see Reimann, 1991). Computer-based simulations suitable for use with this type of learning design typically provide students with the opportunity to set the value of a list of variables or parameters effecting the simulation and then view the effect of these values on the simulated phenomena. For example, Rivers and Vockell (1987), describe a predator-prey simulation in which students can manipulate variables including food supply, carrying capacity and various environmental conditions, and then view the effects through tabular and graphic output.

An alternative learning design approach is to provide students with a specific goal to achieve through their exploration or manipulation of the simulation (over and above the goal of predicting the behavior of the simulation). The assumption here is that if the goal is appropriately aligned to the target learning, achievement of the goal will require students to gain the expected knowledge or understanding. For example Rieber, Tzeng and Tribble (2004) describe a computer-based simulation focusing on Newton’s laws of motion in which participants controlled the movement of a simulated “ball” by applying forces to it in a particular direction with the goal of moving the ball to a target location.

**Support and guidance during discovery learning**

Consistent with the findings from reviews of discovery learning in general such as those undertaken by Mayer (2004) and Alferi et al. (2011) discussed above, empirical studies of discovery learning using computer-based simulations have tended to highlight the need for instructional guidance or support prior to or during the learning process. For example, de Jong and van Joolingen (1998) reviewed a number of studies which compared unsupported discovery learning using computer-based simulations to expository teaching, and found that
the advantages of learning using simulations were not always evident due to the problems learners often encountered with discovery learning. In order to address the problems identified, the authors suggested a range of different types of support and guidance that could be provided for students undertaking these activities, including: the provision of instructional information related to the target learning domain delivered prior to or during exploration; support for hypothesis generation; the provision of experimentation hints during exploration such as advice to change one variable at a time; support for articulating predictions; and the structuring of the learning task and/or the environment characteristics so that the task or environment complexity is initially reduced with the complexity increasing as the learner progresses through the simulation. Examples illustrating some of these support and guidance strategies are provided in the following paragraph.

Rivers and Vockell (1987), in a study in which high school biology students explored a series of computer-based simulations, demonstrate the value of providing instructional information prior to exploring a simulation, including basic background information to the content area as well as advice about strategies for exploring the simulation. Rieber, Tzeng and Tribble (2004) demonstrate the value of providing instructional information during exploration, in their study of the use of a physics simulation in which students were provided with pre-prepared instructional information at intervals during exploration outperformed those without such guidance. An alternative is to provide adaptive system generated information relevant to the learner’s current exploratory activity (see, for example, Leutner, 1993). Njoo and de Jong (1993) describe a study in which in order to help learners be more systematic in their exploration and to focus their attention on important aspects of the simulation, rather than having the students form their own hypothesis, students were provided with a hypothesis and asked to explore its validity. The provision of tasks which gradually increase in complexity is illustrated by immersive environments such as Quest Atlantis.
(Barab et al., 2007) and River City (Ketelhut, Nelson, Clarke, & Dede, 2010). A related approach is to initially limit the degree of learner control, consistent with the findings of Kettanurak, Ramamurthy and Haseman (2001) which suggest that too much learner control can lead to the adoption of ineffective learning strategies or the skipping of important instructional content.

In addition to the support strategies suggested by de Jong and van Joolingen (1998), additional strategies suggested by others include the provision of feedback during goal directed exploratory activities (see, for example, Rieber, Tzeng & Tribble, 2004) and the provision of learning tasks which focus the learner’s attention on particular content or concepts within the simulation. De Jong et al. (1999), demonstrated in a study of students exploring a physics simulation on collision, the way in which assigning specific tasks to learners while exploring a simulation can result in improved learning performance over free exploration (see also Kunsting, Wirth & Paas, 2011, for a study exploring the relationship between the nature of the task goals assigned to students, the learning strategies and learning performance).

**Support through reduced complexity**

An alternative approach to supporting students in discovery learning using a computer-based simulation is to provide students with the output from a series of pre-set simulation parameters which provide them with a systematic exploration sequence rather than asking them to manipulate simulation parameters themselves. Such an approach is consistent with de Jong and van Joolingen’s (1998) idea of reducing the complexity of the simulation and also somewhat similar to Kettanurak, Ramamurthy and Haseman’s (2001) idea of reducing the degree of learner control. With this approach, students are clearly more behaviourally passive in their discovery. However, consistent with Alferia et al.’s (2011) and Piaget’s (1970, 1977) warnings, it is important not to conflate behavioural activity with active (cognitive)
knowledge construction. Clearly, applications of the predict-observe-explain learning design within a computer-based simulation generally involve the learner manipulating the parameters of the simulation prior to making a prediction. However, supporting the discovery learning process by having the system choose the simulation parameters is not completely inconsistent with this learning design and potentially would allow the learner to be guided through a discovery process involving a logical sequence of parameters. The need for, or value of, active manipulation of simulation parameters within the discovery learning process is a question not addressed by earlier studies.

An exploration of the learning consequences of allowing or not allowing the learner to set the parameters within a computer-based simulation has parallels in research comparing learning under learner versus program control within multimedia learning resources. Such research typically has focussed on the value of allowing learners choice in the sequence and selection of content within a learning resource compared, for example, to encountering similar material in a predefined sequence, as they would, for example, when viewing a video or using a traditional computer-based tutorial with a lock-step design (see, for example, Hannafin & Sullivan, 1995). Milheim and Martin (1991) describe three alternative theoretical arguments in support of the learning benefits of learner control: the first drawing on motivation theory with a focus on the way in which learner control can increase the sense of personal relevance of the material; the second drawing on attribution theory which can explain individual differences in the learning benefits from learner control in terms of whether students attribute their success to their own choices and actions; and the third drawing on information processing theories of learning, suggesting that learner control may have a positive effect on the way an individual encodes and stores information and consequently on their schema construction.
Kinzie, Sullivan and Berdel (1988) highlight the theoretical support for learner control and consistent with Milheim and Martin’s (1991) motivational arguments note the clear evidence for attitudinal advantages for learner control, however, also note that the results for learning outcomes are very mixed. Cordova and Lepper’s (1996) finding, that the provision of learner choice in relation to contextual rather than instructional aspects of a learning resource, such as the names of characters within a game, led to greater attention and improved learning outcomes, is also consistent with Milheim and Martin’s (1991) arguments based on motivational theory. When the learner is given control over important instructional decisions, however, similar findings to those emerging from discovery learning research emerge, with some unguided learners making poor control decisions leading to either inefficient exploration of confusion. Kinzie, Sullivan and Berdel (1988), for example, suggest that in order for learning outcome benefits from learner control to emerge, a balance between learner control and instructional guidance is required. Lawless and Brown (1997) concur with this, concluding, in a review of research focussing on the ways in which prior knowledge, self efficacy and interest on the one hand and learner control, instructional design and level of control on the other impact on the learning process, that “while the ability to control one’s instructional sequence can enhance learning and heighten attitudes and self-efficacy, unrestricted control and lack of learning goals can dampen the power of learning in such an environment” (p. 127).

This study
The focus of this study was on the value of learners planning and executing changes to simulation parameters in order to test hypotheses about the behaviour of the simulation. This distinguishes the study from learner-control research which focuses on the value of learners controlling the selection of instructional content to view. The theoretical basis for this investigation was the idea that providing learners with control over the setting of parameters
of a simulation will support their ability to construct and reconstruct cognitive schema (rather than the idea that such control will lead to attentional or motivational benefits).

Specifically, this study compared the discovery learning performance of students actively exploring a simulation by setting their own parameters (referred to as the ‘Exploration’ condition) with the performance of students viewing the simulation output from a sequence of preset parameters (referred to as the ‘Observation’ condition). A finding in favour of the Exploration condition would be consistent with arguments for the value of ‘active’ learning processes, drawing on Piaget’s notion of cognitive schema construction. A finding in favour of the Observation condition would support the arguments against unguided discovery learning and particularly the arguments that guidance or reduced complexity is needed if discovery learning is to be effective, consistent with the ideas of Kirschner, Sweller and Clark (2006), Mayer (2004), and Alferi et al. (2011).

**Methods**

**Design**

As shown in Table 1, the study compared the learning performance of Exploration participants and Observation participants using computer-based simulations in two content areas, Blood Alcohol Concentration and Global Warming. The study was conducted with 158 University of Wollongong education students. Each student completed the Exploration condition using a computer-based simulation within one content area and the Observation condition using a simulation within the other content area. The order was randomised so that approximately half of the students completed the Observation condition first while the other half completed the Exploration condition first. Prior to each learning condition the students completed a knowledge pretest. After each learning condition the students completed a
knowledge posttest including the same items as the pretest. Students also completed some additional open ended test items in the posttest along with an engagement questionnaire, the results from which will be the focus of other publications and are not discussed in this article. Student actions within the learning resources were logged by the software to allow later analysis of their exploration strategies.

For analysis purposes the activities within the two content areas were treated as distinct experiments. In each case the independent variable was learning condition, with two levels, Exploration and Observation, and the dependent variable was learning performance. Undertaking two concurrent experiments using two distinct content areas with students undertaking one learning condition in one content area and the other learning condition in the other content area (see, for example, Lindström et al., 1993) provided us with two distinct data sets to draw upon in making conclusions from the data, which we considered important, given the potential for content or resource specific factors to impact on the results.

Pilot testing of the simulations and the experimental protocols, involving user observation and follow up interviews was undertaken in two phases, each involving two pilot testers, with changes made to the user-interface and to the screen designs as a result of each phase. This allowed us to be confident that the instructions provided to students were clear and that learning activities using the computer-based simulations would allow learning of the target concepts.

Learning conditions

The content areas of Blood Alcohol Concentration and Global Warming were chosen, firstly because they were thought to be of interest to members of the participant community, and secondly because each contained common misconceptions which could potentially be
uncovered and addressed through exploration of a computer-based simulation. The computer-based simulations used in the study each contained a conceptual model illustrating key concepts within the content area (see de Jong & van Joolingen, 1998 for a discussion about the distinction between simulations containing conceptual versus operational models). The versions used in the exploration conditions were designed to allow students to follow a ‘predict-observe-explain’ learning design pattern (White & Gunstone, 1992), with students encouraged to mentally predict the effect of their simulation parameter changes, observe the results of the change and mentally try to explain the observed results. Figures 1 and 2 provide excerpts from the simulations used in the Exploration conditions showing the screens on which participants could change variables on the left-hand side of the screen before running the simulation to see the results. In the case of global warming, the four graphs on the right-hand side of the screen, showing values for ozone layer thickness, CO₂ concentration, greenhouse insulation and surface temperature, changed if the variables were altered. In the blood alcohol concentration simulation the graph on the right-hand side showing values for blood alcohol concentration over time, changed if the input variables were altered. As far as possible the user interface was consistent across the two topics to maximise comparability between the resources.

INSERT FIGURE 1 HERE

INSERT FIGURE 2 HERE

The simulations used in the Observation condition for each topic area consisted of a series of screens containing pre-defined values on the left hand side and the relevant graphical outputs on the right hand side. Students were provided with a single option, a continue button, which resulted in the next set of pre-defined values being loaded and the corresponding output displayed. This is illustrated in Figures 3 and 4. The pre-defined values were chosen to
demonstrate the effect of each parameter on the simulation output. In most cases the parameters chosen contained either one parameter different from the benchmark values shown (2006 values in the case of global warming or ‘Bill’s values’ in the case of blood alcohol concentration) or one parameter different from the parameters on the previous screen.

INSERT FIGURE 3 HERE

INSERT FIGURE 4 HERE

Similar to de Jong et al. (1999) and as suggested by Rivers and Vockell (1987) we included explanatory material at the beginning of each resource designed to provide definitions of key terms and sufficient background to the content areas to ensure that students could explore without confusion. Students progressed through this material at their own pace, one screen at a time before commencing their iterations through either the Exploration or Observation screens shown above. As discussed by de Jong and van Joolongen (1998), insufficient domain knowledge can be a key contributor to poor outcomes from unguided discovery learning.

**Pretests and posttests**

The knowledge pretests and posttests contained a series of identical items designed to measure understanding of key concepts within the content area. For example, within the Blood Alcohol Concentration content area, the following item was included within the pretest and posttest to explore the students’ understanding of the relationship between of a person’s weight and their blood alcohol concentration:

*A person with greater body weight:*

a) *Will have a higher Blood Alcohol Concentration (BAC) than a lighter person.*

b) *Will have a lower BAC than a lighter person.*
c) Will have the same BAC as a lighter person.

d) Will have their BAC increase at a greater rate than a lighter person

The following item was included within the Global Warming pretest and posttest to test the students on their understanding of the way various environmental factors affect the global temperature:

*Which of the following environmental factors have a direct or an indirect effect on the Global Average Surface Temperature (GAST)?*

a) The amount of Carbon Dioxide (CO2) absorbed by plants

b) The thickness of the ozone layer

c) The percentage of CO2 in the atmosphere

d) The Greenhouse insulation effect

Some questions were specifically designed to test whether students held certain common misconceptions within each content area. Some of the questions, such as the Global Warming example question above, had multiple valid stems while some, such as the Blood Alcohol Concentration example question above, had a single valid stem. Students scored one point for each question in which only the valid stems were chosen, that is, in order to score a point on a question their answer had to be completely correct – no partial marks were awarded. There were seven questions in the Global Warming test and nine questions in the Blood Alcohol Concentration test.

**Results**

The first analysis considered whether Observation or Exploration participants showed improved understanding from pretest to posttest as a result of their learning condition. An initial analysis of the learning which occurred in each condition was undertaken using paired *t* tests which compared the pretest and posttest scores (Table 2). The results of this analysis
indicated that for the Global Warming content area, there was a slight decrease in test performance from the pretest to the posttest for Observation participants and no difference in performance for Exploration participants. For the Blood Alcohol Concentration content area there was no difference in test performance for the Observation participants and a slight increase for Exploration participants.

For each content area (Global Warming and Blood Alcohol Concentration) an analysis of covariance (ANCOVA) test was completed which included posttest score as the dependent variable, learning condition as the independent variable (Exploration, Observation), and pretest score as a covariate. By using the pretest score as a covariate in the comparison of student posttest scores, students’ learning for the different learning conditions within each content area was able to be compared, regardless of prior knowledge. As shown in Table 3, for the Global Warming content area there was no effect for learning condition (F (1,155) = 2.40; p = .124). For the Blood Alcohol Concentration, there was a significant effect for condition (F (1, 155) = 5.52; p = .02), with Exploration participants showing a significant increase in understanding compared to Observation participants. There was a significant covariant effect for students’ pretest scores on their posttest understanding in both the global warming (F (1,155) = 27.50; p < .001) and the blood alcohol concentration (F (1,155) = 16.40; p < .001) content areas. This is unsurprising but supports the inclusion of pretest score as a covariant in the analysis.

As can be seen from Table 2, despite students showing significant gains in performance in the Blood Alcohol Concentration content area, in general students showed very little improvement from pretest to posttest, regardless of content area or condition. Given the
maximum scores that could be obtained by students, performance scores were well below the midpoint of the scale for both content areas. The lack of substantial improvement from the pretest to the posttest by many students was somewhat surprising. In order to understand this more fully, a closer analysis of the variation in students’ performance scores was undertaken. It was noted that the variance of students’ posttest scores in the Exploration condition, particularly for the global warming content area, was quite high, possibly reflecting differences in how the simulation was used and explored by students. A preliminary analysis showed that some students seemed to have explored the simulation systematically and others were more unsystematic in their approach. This warranted more structured investigation.

An analysis was undertaken to develop a greater understanding of whether some students manipulated simulation parameters in ways which were likely to lead to greater understanding. Two strategies by which a student could systematically explore the simulation were identified. One strategy would be for students to work their way through the simulation changing one variable in each iteration, comparing the output from each iteration with their recall of the output from the previous one. An alternative strategy would be for students to make use of the preset values provided as a comparison point on each output screen (‘2006 values’ in the case of the Global Warming simulation and ‘Bill’s values’ in the case of the Blood Alcohol Concentration simulation) and manipulate the parameters so that in each iteration values matching these preset values were used for all except one variable.

Table 4 provides a summary of the frequency with which students adopted each of these strategies for a given number of iterations within the simulator. Table 4 also shows the mean posttest scores for students within each of these categories. This analysis suggested that a large number of students did not at any stage in their exploration change only one variable from the previous iteration (34 students in global warming and 21 students in blood alcohol concentration) and a large number did not at any stage ensure that only one variable was
different from the provided values (35 students in global warming and 38 students in blood alcohol concentration). Those who did change one variable at a time or vary one variable from the provided values in a number of their iterations through the simulation tended to perform better in the posttest.

In order to further explore this apparent finding Exploration participants’ strategies were classified either as systematic (frequently changing one variable at a time), or unsystematic (in general changing multiple variables at a time). Specifically, we characterised Exploration participants’ interaction strategies as Systematic Exploration if they explored the simulation by changing only one variable from the provided example (e.g. ‘Bill’s values’ or ‘2006 values’) or from previous simulation outputs on four or more occasions. All other Exploration participants’ interaction strategies were classified as Unsystematic Exploration. The threshold value of four was chosen because it was regarded as the minimum number of systematic iterations necessary to allow learning of a number of the key concepts, and this threshold also ensured was that we had sufficiently sized sub-groups.

INSERT TABLE 4 HERE

Another series of inferential tests were completed which mirror those reported above except that the independent variable (condition) was now a three level variable (Observation, Unsystematic Exploration and Systematic Exploration). Table 5 shows the comparison of pretest and posttest scores for participants in each of these conditions. These tests indicated that the difference between pretest and posttest scores were significant for the Blood Alcohol Systematic Exploration participants but not for Global Warming Systematic Exploration participants, while there was no difference between pretest and posttest scores for Unsystematic Exploration participants in either content area.

INSERT TABLE 5 HERE
An ANCOVA for each content area was calculated with a three-level independent variable (Observation, Unsystematic Exploration, and Systematic Exploration) and pretest score as the covariate. As shown in Table 6, significant main effects were recorded for both content areas (and there were significant covariate effects for both Global Warming (F (1,154) = 23.72; p < .001) and Blood Alcohol (F (1,154) = 23.72; p < .001)). Orthogonal contrasts indicated that students in the Systematic Exploration group recorded significantly higher posttest scores than students in both the Unsystematic Exploration and the Observation groups and there were no differences between posttest scores for these latter two groups. This pattern of results was the same for both content areas.

**INSERT TABLE 6 HERE**

**Discussion**

The comparison of the posttest performance of Exploration participants who actively explored and manipulated a simulation with Observation participants who observed simulation output from preset parameters found a small but significant difference in test performance in one content domain and no significant difference in the other. Our attempts to understand this result led us to scrutinise the strategies used by Exploration participants by examining the log files which recorded their interactions during exploration. Our examination of these log files in conjunction with posttest results suggested that some participants appeared to have explored the simulations quite unsystematically and that these poor exploration strategies may have contributed to poor posttest performance for these students. The fact that some participants explored the simulations systematically and some did not is consistent with early observations by Bruner (1961) about the way students go about the classic 20 questions game, and also with the findings from a number of other discovery learning studies (see, for example Glaser et al., 1991; Reimann, 1991; Kuhn et al, 1992).
Our analysis of the interaction log files allowed us to identify a number of variables that we could have used to characterise students’ strategies, including time spent on each screen, number of iterations through the simulation, the number of variables changed during each simulation, and the actual values used. Similar to Kennedy and Judd (2004) who used cluster analysis to identify clusters of students with interaction patterns illustrating distinct learning strategies while exploring a digital learning resource, we experimented with cluster analysis to see whether clusters of students with similar strategies would emerge. Although this avenue was seen as promising and may be used in future analysis, we eventually settled on a criteria which focused on the key aspect of the students’ exploration strategies which we felt allowed us to differentiate between systematic and unsystematic explorers, namely the number of variables changed or varied during each iteration. Another alternative we could have considered and may explore in future analysis is a characterisation based on the successive choice of values for parameters. For example, Thompson and Reimann (2010), drawing on Levy and Wilensky (2005), used rules based on the values chosen by learners, the time spent and the number of iteration, and characterised learner strategies as ‘straight to the point’, ‘homing’ or ‘oscillating’, in manipulating an agent-based model.

Once students’ exploration strategies were characterised as either systematic or unsystematic, we were able to compare posttest performance of participants with these different strategies. This comparison showed that Systematic Exploration participants performed better than Unsystematic Exploration participants in both content areas. Additionally, Systematic Exploration participants in each content area performed significantly better than Observation participants but there was no difference between the performance of Observation participants and Unsystematic Exploration participants. This suggests that in unguided discovery learning, for some students, the freedom to manipulate the simulation and explore the effects of chosen parameter values on the simulation output
can be effective. For other students, however, the benefits of active exploration are countered by the confusion caused by unsystematic exploration strategies and consequently they perform no better than they would perform after passively viewing the simulation output from a series of preset parameter values.

The key implication of these results for designers of computer-based instructional simulations is the need for guidance during exploration which identifies students exploring unsystematically and reminds them of the need to explore systematically and in particular the need to change one parameter at a time. This is consistent with the recommendations of de Jong and van Joolingen (1998). In situations where such guidance during exploration is not feasible, it may be that the constraint provided by having the parameter values set for students could be a valuable form of scaffolding which reduces the complexity of the task, consistent with the complexity reducing approaches adopted by Barab et al. (2007), Ketelhut et al. (2010) and Kettanurak et al. (2001).

The fact that students who explored the simulation systematically outperformed students who observed simulation output but did not have the opportunity to set simulation parameters provides support for constructivist theories of learning. A constructivist interpretation of these results would be that the opportunity to establish a personal hypothesis about the relationship between simulation parameters and simulation output and to set the parameters in a way which would allow this hypothesis to be tested, better supported the students in an equilibration process involving iterative construction and reconstruction of a personal representation of knowledge within the content area of the simulation. The fact that many students did not appear to benefit from this opportunity is consistent with the findings of earlier studies summarised by Mayer (2004) and Alferi et al. (2011). Additional analysis of these earlier results is needed to determine whether unsystematic exploration could explain or partially explain the modest learning performance of unguided discovery learners during
these earlier studies, or whether Kirschner, Sweller and Clark’s (2006) explanation focusing on the inherent cognitive load associated with discovery learning would be more accurate.

In interpreting these results an important question that needs to be asked is: what led some students to explore the simulation systematically and others unsystematically? Kirschner, Sweller and Clarke (2006) suggest that the value of and the need for explicit instructional guidance only begins to diminish when students have sufficient prior knowledge to allow for internal guidance. Schauble et al. (1991) also demonstrate the way in which the level of sophistication in students’ mental models prior to learning can have an impact on the level of sophistication in their learning strategies during a discovery task. These ideas would suggest that the reason some students explored systematically was that they had sufficient prior knowledge to allow them to avoid becoming confused or overwhelmed by the exploratory process. The fact that Systematic Exploratory participants had noticeably higher pretest scores than Unsystematic Exploratory participants in both content areas is consistent with this idea (see Table 6). A $t$ test comparing systematic versus unsystematic participants’ pretest scores found that the differences were not significant in the blood alcohol content area ($t(83) = 1.185, p=0.239$) but were significant in the global warming content area ($t(71)=2.107, p=0.039$). Further research is needed which explores the degree to which students’ level of prior knowledge affects their exploration strategy in discovery learning activities.

An alternative explanation for why some students explored systematically and some did not is that some students may have a natural aptitude for systematic exploration or through prior experience of more structured inquiry-based learning designs, such as problem-based learning (see Hmelo-Silver, Duncan & Chinn, 2007), they have developed more systematic exploratory approaches. This idea is consistent with some of the early findings of Anthony (1973) demonstrating the link between success during a discovery learning activity and subsequent ability to apply discovery approaches in other contexts.
In general the learning that occurred as a result of the various learning conditions was modest and this was particularly the case for the global warming content area where neither students undertaking the Observation condition nor students undertaking the Exploration condition achieved significantly better on the posttest than the pretest. This modest learning performance to some extent reflects the complexity of the material. The students undertaking the study were first year university students training to be primary school teachers and anecdotal evidence suggests that the proficiency of these students in scientific reasoning is quite varied. Additionally, although the content areas were intended to be of interest to the student population, the fact that learning of the principles within the two content areas was not specifically aligned to the students’ area of study, may have impacted on their motivation for the task. De Jong and van Joolingen (1998) also noted the difficulty many students experience in interpreting graphical output from simulations, and it is likely that the output from the global warming simulation, which consists of four separate graphs, was more difficult to interpret than the output from the blood alcohol simulation which appeared as a single graph.

The scoring system used required students to identify each valid stem and to not incorrectly identify any invalid stems on each question in order to get a mark. This meant that in some cases a student may have partially understood a concept but not been awarded a mark because of one misconception. This scoring system was used because it ensured that we could measure students’ complete understanding of the key concepts within each content area. The fact that partial marks were not awarded where some stems were correct may, however, have also contributed to the low raw scores obtained by many students. Nevertheless, the lack of substantial improvement from the pretest to the posttest by many students was somewhat surprising. To some extent this result supports the arguments of Mayer (2004), Kirschner, Sweller and Clark (2006) and Alferi et al. (2011), who suggest that
pure, unguided, discovery learning is unlikely to be effective. Although we implemented some recommendations from de Jong and van Joolingen (1998) regarding the types of guidance needed in discovery learning activities using computer-based simulations to ensure that they are effective (for example we included instructional material in advance of the discovery activity), the absence of support during the task, such as experimentation hints or opportunities to explicitly articulate predictions, may have contributed to the poor learning performance of some students.

The fact that some students performed worse in the global warming posttest than in the pretest led us to scrutinise the results to see whether there was any consistency in the questions answered incorrectly after the learning conditions, with a view to identifying any misconceptions emerging through the use of the simulation or any ambiguities in the test questions. This scrutiny led us to conclude that there was no consistency to the degradation in performance from pretest to posttest and no obvious problems with either the simulation or the test questions. Rather, it would appear that the complexity of the material led some students to become more confused and consequently to be less sure about aspects of the content area that they previously had understood correctly.

**Conclusions**

This article has reported on a study focusing on the performance of students undertaking discovery learning using computer-based simulations and in particular comparing the performance of students given the opportunity to explore and manipulate simulations using a predict-observe-explain approach with the performance of students who were constrained to observing simulation output from a series of preset parameter values. The results suggest that without guidance during exploration the opportunity to freely explore and manipulate the simulation can lead to confusion and poor learning performance for some students. However, for other students, this active exploration process can be more effective than passive
observation. A key recommendation for designers of computer-based instructional simulations is that guidance is required that reminds students during their exploration of the need to explore the simulation systematically and in particular to vary one parameter at a time.

The results provide further support to the arguments of Mayer (2004), Kirschner, Sweller and Clark (2006) and Alferi et al. (2011), who suggest that pure, unguided discovery learning can be ineffective for many learners. The results also provide support for the arguments of de Jong and van Joolongen (1998) who discuss in detail the types of guidance needed by learners undertaking discovery learning activities using computer-based simulations. Finally, the results add to understanding about the value of active exploration during discovery learning, suggesting that although restrictions to the learner’s ability to manipulate the simulation can provide a valuable scaffold in some instances, if sufficient guidance is provided during exploration it would be reasonable to conclude that active exploration could lead to improved learning outcomes over the alternative of passive observation of simulation output.

References


<table>
<thead>
<tr>
<th>Condition</th>
<th>Blood Alcohol Concentration</th>
<th>Global Warming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation</td>
<td>N=73</td>
<td>N=85</td>
</tr>
<tr>
<td>Exploration</td>
<td>N=85</td>
<td>N=73</td>
</tr>
</tbody>
</table>

Table 1. Experimental design.
Table 2. Mean pretest and posttest scores for both conditions across the two content domains

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Condition</th>
<th>Pretest M (SD)</th>
<th>Posttest M (SD)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Warming+</td>
<td>Observation (n=85)</td>
<td>1.82 (1.51)</td>
<td>1.42 (1.29)</td>
<td>2.26</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>Exploration (n=73)</td>
<td>1.68 (1.42)</td>
<td>1.72 (1.85)</td>
<td>0.20</td>
<td>0.841</td>
</tr>
<tr>
<td>Blood Alcohol^</td>
<td>Observation (n=73)</td>
<td>3.55 (1.25)</td>
<td>3.42 (1.31)</td>
<td>0.60</td>
<td>0.552</td>
</tr>
<tr>
<td></td>
<td>Exploration (n=85)</td>
<td>3.60 (1.24)</td>
<td>3.93 (1.40)</td>
<td>2.33</td>
<td>0.022</td>
</tr>
</tbody>
</table>

^max = 9; +max = 7
Table 3. ANCOVA comparisons of mean posttest scores between conditions for each content area.

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Observation M (SD)</th>
<th>Exploration M (SD)</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Warming</td>
<td>1.42 (1.29)</td>
<td>1.72 (1.85)</td>
<td>2.40</td>
<td>0.124</td>
</tr>
<tr>
<td>Blood Alcohol</td>
<td>3.42 (1.31)</td>
<td>3.93 (1.40)</td>
<td>5.52</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Table 4. Frequency of students exhibiting systematic strategies for each number of iterations within each content area and mean posttest scores.

<table>
<thead>
<tr>
<th>Content Area and Strategy</th>
<th>Number of Iterations</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7+</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Global Warming</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One variable different from previous</td>
<td>F (M)</td>
<td>F (M)</td>
<td>F (M)</td>
<td>F (M)</td>
<td>F (M)</td>
<td>F (M)</td>
<td>F (M)</td>
<td>F (M)</td>
<td>F (M)</td>
</tr>
<tr>
<td>One variable different to provided values</td>
<td>34</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Mean</td>
<td>1.3</td>
<td>1.6</td>
<td>1.8</td>
<td>1.0</td>
<td>2.6</td>
<td>2.6</td>
<td>2.0</td>
<td>2.2</td>
<td>2.1</td>
</tr>
<tr>
<td>One variable different from previous</td>
<td>35</td>
<td>18</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Mean</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>3.0</td>
<td>1.6</td>
<td>2.5</td>
<td>3.5</td>
<td>2.6</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>Blood Alcohol Concentration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One variable different from previous</td>
<td>21</td>
<td>16</td>
<td>12</td>
<td>9</td>
<td>7</td>
<td>9</td>
<td>4</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Mean</td>
<td>3.1</td>
<td>3.8</td>
<td>3.5</td>
<td>4.7</td>
<td>4.1</td>
<td>4.4</td>
<td>4.0</td>
<td>5.0</td>
<td>4.2</td>
</tr>
<tr>
<td>One variable different to provided values</td>
<td>38</td>
<td>17</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>3.2</td>
<td>3.7</td>
<td>4.2</td>
<td>5.6</td>
<td>NA</td>
<td>4.3</td>
<td>5.3</td>
<td>4.9</td>
<td>3.9</td>
</tr>
</tbody>
</table>
Table 5. Mean pretest and posttest scores for participants in each condition for each content area.

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Condition</th>
<th>Pretest M (SD)</th>
<th>Posttest M (SD)</th>
<th>T</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Warming+</td>
<td>Observation (n=85)</td>
<td>1.82 (1.51)</td>
<td>1.42 (1.29)</td>
<td>2.26</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>Unsystematic Exploration (n=48)</td>
<td>1.44 (1.18)</td>
<td>1.33 (1.52)</td>
<td>0.44</td>
<td>0.662</td>
</tr>
<tr>
<td></td>
<td>Systematic Exploration (n=25)</td>
<td>2.16 (1.72)</td>
<td>2.48 (2.20)</td>
<td>0.83</td>
<td>0.415</td>
</tr>
<tr>
<td>Blood Alcohol^</td>
<td>Observation (n=73)</td>
<td>3.55 (1.25)</td>
<td>3.42 (1.31)</td>
<td>0.60</td>
<td>0.552</td>
</tr>
<tr>
<td></td>
<td>Unsystematic Exploration (n=51)</td>
<td>3.47 (0.17)</td>
<td>3.51 (1.30)</td>
<td>0.24</td>
<td>0.814</td>
</tr>
<tr>
<td></td>
<td>Systematic Exploration (n=34)</td>
<td>3.79 (1.25)</td>
<td>4.56 (1.33)</td>
<td>3.25</td>
<td>0.003</td>
</tr>
</tbody>
</table>

+max = 7; ^max = 9
Table 6. ANCOVA comparisons of mean posttest scores between conditions (three levels) for each content area.

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Observation M (SD)</th>
<th>Unsystematic Exploration M (SD)</th>
<th>Systematic Exploration M (SD)</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Warming</td>
<td>1.42 (1.29) a</td>
<td>1.33 (1.52) a</td>
<td>2.48 (2.20) b</td>
<td>4.17</td>
<td>.017</td>
</tr>
<tr>
<td>Blood Alcohol</td>
<td>3.42 (1.31) a</td>
<td>3.51 (1.30) a</td>
<td>4.56 (1.33) b</td>
<td>8.69</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

^ab different superscript across rows indicate between group differences (p < .001).
Figure 1. Excerpt from the Global Warming Exploration Simulation.

Figure 2. Excerpt from the Blood Alcohol Concentration Exploration Simulation.

Figure 3. Excerpt from the Global Warming Observation Simulation.

Figure 4. Excerpt from the Blood Alcohol Concentration Observation Simulation.