Optimizing Learning of Scientific Category Knowledge in the Classroom: The Case of Plant Identification

By: Bruce K. Kirchoff, Peter F. Delaney, Meg Horton, Rebecca Dellinger-Johnston

Kirchoff, B. K., P. Delaney, and M. Horton. (2014). Optimizing Learning of Scientific Category Knowledge in the Classroom: The Case of Plant Identification. *CBE Life Sciences Education, 13(3)*, 425-436. doi: 10.1187/cbe.13-11-0224

Keywords: Plant Systematics | Plant Identification | Learning

CBE—Life Sciences Education
Vol. 13, 425–436, Fall 2014

Article

Optimizing Learning of Scientific Category Knowledge in the Classroom: The Case of Plant Identification

Bruce K. Kirchoff,* Peter F. Delaney,† Meg Horton,‡
and Rebecca Dellinger-Johnston*

*Department of Biology, †Department of Psychology, and ‡Grogan College, University of North Carolina at Greensboro, Greensboro, NC 27402-6170

Submitted November 24, 2013; Revised June 9, 2014; Accepted June 10, 2014
Monitoring Editor: Diane Ebert-May

Learning to identify organisms is extraordinarily difficult, yet trained field biologists can quickly and easily identify organisms at a glance. They do this without recourse to the use of traditional characters or identification devices. Achieving this type of recognition accuracy is a goal of many courses in plant systematics. Teaching plant identification is difficult because of variability in the plants' appearance, the difficulty of bringing them into the classroom, and the difficulty of taking students into the field. To solve these problems, we developed and tested a cognitive psychology–based computer program to teach plant identification. The program incorporates presentation of plant images in a homework-based, active-learning format that was developed to stimulate expert-level visual recognition. A controlled experimental test using a within-subject design was performed against traditional study methods in the context of a college course in plant systematics. Use of the program resulted in an 8–25% statistically significant improvement in final exam scores, depending on the type of identification question used (living plants, photographs, written descriptions). The software demonstrates how the use of routines to train perceptual expertise, interleaved examples, spaced repetition, and retrieval practice can be used to train identification of complex and highly variable objects.

DOI: 10.1187/cbe.13-11-0224
Address correspondence to: Bruce K. Kirchoff (kirchoff@uncg.edu).

Conflict of interest statement: B.K.K. is the joint owner of Metis LLC, the company that produces the software used in these studies. This software is distributed free of charge on the company website: www.metisllc.com. The main experiment reported here was designed by M.H., carried out by B.K.K., and analyzed by P.F.D. R.D.-J. designed, conducted, and assisted in the analysis of the retest, under the direction of B.K.K. and P.F.D. The separation of these functions helped ensure that B.K.K.'s apparent conflict of interest is not reflected in our published results. No promotion of the products of Metis LLC to the exclusion of other similar products should be construed from his role as an author of this paper.

INTRODUCTION

Scientific expertise in many disciplines depends on learning to accurately classify novel (unstudied) examples into categories. In plant systematics, an expert can rapidly classify novel individuals seen in nature into the correct species—a process known as plant identification. None of the examples seen in the field will be exactly the same as the ones studied during training, and so the classification task involves generalization. It involves the correct assignment of representative organisms or images to pre-existing categories, the taxa to which they belong.

In this paper, plant identification serves as a model for testing the efficacy of applying recently explored principles in the cognitive psychology of learning and memory to a natural science classroom. The software used for this purpose incorporates several methods of improving learning that have been validated in psychology research: 1) ideas from perceptual learning (holistic visual processing); 2) the use of quizzes and tests to increase learning via memory retrieval (testing

Supplemental Material can be found at:
http://www.lifescied.org/content/suppl/2014/08/29/13.3.425.DC1.html

effects/retrieval practice); 3) interleaved practice, which involves mixing of different taxa within the same study session, as opposed to blocked learning, in which all the examples of one taxon are presented together; and 4) review of practiced materials multiple times after their initial presentation, with time between repetitions (spaced practice). We review the scientific support for each of these principles below when we describe our cognitive analysis of plant identification.

These techniques can be applied to many other learning domains, some of which will be described in more detail in the discussion. For instance, the ability to distinguish species of snakes, lizards, and skinks is important for herpetologists, but it is difficult to learn due to the presence of heteromorphic characteristics that are often present in animals of the same species. Limited exposure to the organisms in the classroom makes it difficult to generalize to novel examples when they are encountered in the field. In organic chemistry, the ability to classify novel members of a family of molecules (e.g., esters) from their chemical structure requires the chemist to orient on the relevant part of the structure—which is invariant for all members of a family—and ignore the parts that change from example to example. Learning which parts of the molecule to attend to is the critical problem and is a skill that is not currently actively taught.

The Challenges of Plant Identification

A core goal of plant systematics, field botany, horticulture, and many agriculture courses involves learning to identify plants by categorizing them into an established taxonomy. A biological taxonomy is a named, hierarchical arrangement of organisms. A taxon (plural taxa) is a group of organisms at one of the levels of the hierarchy. The levels of hierarchy that are of most concern to beginning students are the family, genus, and species. The higher-level categories (order, class, division, etc.) are more relevant to advanced students and will not be discussed in this paper. Species are the most basic level of the taxonomic hierarchy. They are grouped together into genera (singular genus), which in turn are grouped into families. For example, all of the species of hickory trees, taken as a unit, form the genus *Carya*. Walnut trees collectively form the genus *Juglans*. The genera *Carya* and *Juglans* are members of the family Juglandaceae (the walnut family).

Learning to identify living things is quite difficult. The examples of *Carya* and *Juglans* illustrate some of what makes the task so hard: the genera are visually foreign to students, are named in a foreign language, are difficult to observe in nature, and embody a good deal of variability. The bark of *Carya ovata* (shagbark hickory) does not closely resemble that of *Carya cordiformis* (smooth bark hickory), yet both belong to the same genus and must be recognized as such. Plant identification thus requires a great deal of study to master. Of all organisms, plants are perhaps the most difficult, not least for the reason that most undergraduates have little or no previous exposure to them (Wandersee and Schussler, 1999; Schussler *et al*., 2010; Stagg and Donkin, 2013) and thus have little interest in learning their names. Perhaps because of this lack, images of plants are harder to recall than are images of animals (Schussler and Olzak, 2008). To make matters worse, identifying an unknown plant when it is not immediately recognized can be time-consuming due to the number of choices that must be made when using botanical keys (Tilling, 1984) and the difficulty of the characters used in the identification (Fermanian *et al*., 1989). Knowing the family or genus of the unknown can greatly reduce both the time required and the likelihood of error by reducing the number of characters that must be correctly described during the keying process. Thus, the development of visual expertise has long been an important learning objective for botanists.

One way of acquiring expertise in plant recognition involves repeated exposure to the plants in a natural setting. By repeatedly seeing the living plants identified by an expert and being repeatedly quizzed on them, the student gains proficiency in species recognition. However, this method is effective only if the student supplements in-class exposure with effective study outside of class. Class time is not sufficient to learn reliable identification. Exposure to a range of variation within a species is also necessary. Concepts , including species concepts, contain representations of both typical examples of the category and embody information about variation within the category. Both of these factors are necessary to form adequate concepts (Fried and Holyoak, 1984; Medin, 1989; Wisniewski, 2002; Perlman *et al*., 2012).

Although field botany courses are effective at teaching plant identification, they are both labor and time intensive, and have been falling out of use at larger universities. They are particularly problematic in urban universities such as ours, where we do not have easy access to natural settings and have a relatively large number (48) of students enrolled in our plant systematics course each year. The software described and tested in this paper was developed by the first author partially as a substitute for more traditional, field-based pedagogy, and as a method of assuring that students used the most effective study methods outside the classroom.

In the next few sections, we outline our cognitive analysis of the problem of plant identification and explain the reasons why we expect the software to be effective. We then describe the program in more detail and go on to report an experimental test of its efficacy in the classroom.

Cognitive Analysis of the Problem of Plant Identification

One of the difficulties in applying cognitive psychology to pedagogical problems is that there are a variety of remedies that could be used to address a specific problem. It requires some expertise to analyze a task and recognize which kinds of interventions are most likely to be effective. At the same time, it requires domain expertise to know which types of competencies are critical to the learning objectives of the course. Plant identification serves as an interesting case study in this regard.

Plant Identification as Holistic Perceptual Expertise. Plants present complex visual images that contain multiple defining features. Many of these features overlap or are similar among different taxa, which is one reason why keys are so difficult to use. Because of these characteristics, plant identification classes usually focus on helping students learn defining features and how to apply these features to taxon identification. However, research in visual perception suggests that for complex stimuli with many features, perceptual expertise develops holistically, not by parts. For example, Gauthier and Tarr (1997) trained participants to identify fictional animals called Greebles. The authors found that participants exhibited changes in visual perception that were achieved more quickly and were better retained when the entire object, rather than individual features, was studied, suggesting a whole-object advantage. Extending this work, Gauthier *et al.* (1999) found that training people to recognize Greebles recruits the fusiform face area—the same area used to support face recognition, a perceptual skill that has been shown to be holistic (Young *et al.*, 1987; Hole, 1994; Farah *et al.*, 1998). Gauthier *et al.* (2000) found that perceptual experts in car recognition and perceptual experts in ornithology similarly use the fusiform face area. Taken together, these results suggest that becoming a perceptual expert is holistic and that many types of complex objects are processed holistically, in the same way as faces (Bukach *et al.*, 2006). As the ultimate goal of perceptual expertise is holistic, learning parts as a starting point is likely counterproductive for the final goal. For these reasons, the software tested here does not point out specific features that define the categories. Instead, perceptual learning is developed holistically by rapidly presenting whole images. This type of training forms the basis for rapid field identification, which is what field botanists, horticulturalists, and agronomists are trained to do.

Plant Identification as Categorization. Aspects of plant identification are structurally identical to studies on category learning. A given example of a plant needs to be categorized into its species, even though each example is unique and therefore may contain only some of the features that define the category. Categorization research is a branch of cognitive psychology that has investigated this phenomenon. Recent laboratory studies of category learning have demonstrated that interleaving different examples rather than blocking of a single example type produces superior categorization performance on novel examples. For example, Kornell and Bjork (2008) used paintings in the style of particular artists and asked people to learn to classify/identify the paintings by artist. They found that interleaved study of the examples of artists' works greatly improved the ability to identify novel paintings compared with blocked study of one artist's work, followed by blocked study of another artist's work. This result has been replicated frequently both with artists (Vlach *et al.*, 2008; Kang and Pashler, 2012) and with other items, such as differentiating species of penguins (Kornell *et al.*, 2010; Wahlheim *et al.*, 2011). However, to our knowledge, the results of these studies have never been applied to category learning in the classroom.

Plant Identification as Cued Recall of a Foreign Vocabulary. Identifying taxonomic categories requires not only correct categorization but also retrieval of difficult Latin names. Tasks in which a cue (such as a picture of a plant) is presented and followed by the requirement that the subject recalls the associated target (such as the name of the plant) are called cued-recall tasks. Learning plant identification is therefore similar to other cued-recall tasks that have been studied, such as learning foreign vocabulary words. Manipulations that enhance cued recall should also be effective in enhancing learning of plant names. Among these manipulations are spacing and testing (Delaney *et al.*, 2010; Carpenter *et al.*, 2012). Many instructors naturally include exercises that have spacing or testing components, including having students answer questions in class, administering quizzes, using flash cards, and repeating the same material on different days. However, deliberately maximizing the use of spacing and testing is a distinct advantage of the Visual Learning—Plant Identification (VL-PI) program.

Spacing of practice can occur both within a session, when the same item is studied multiple times with gaps between the presentations, or between sessions, when the same material is repeated on different days. Spacing has been shown to enhance learning of face–name pairs (Carpenter and DeLosh, 2005) and of vocabulary, even over periods of several years (Bahrick, 1979; Bahrick and Phelps, 1987; Bahrick *et al.*, 1993; Dempster, 1987; Sobel *et al.*, 2011). If the intervals between repetitions of the same item are filled with other study items, spacing also naturally implements the interleaving of exemplars from different categories that we described earlier.

Active Learning. Requiring people to answer test questions is an exercise in active, versus passive, learning. In active learning, the student interacts with the material that is to be learned in some way, rather than reading (text) or viewing it (images, videos) without interaction. Active learning of this sort can take place either inside or outside the classroom and has been shown to aid memory and learning both in laboratory settings (Roediger and Karpicke, 2006; Delaney *et al.*, 2010; Karpicke and Blunt, 2011) and in science classrooms (McDaniel *et al.*, 2011). Active learning has memory benefits even years after the initial learning took place (Pashler *et al.*, 2007). The benefits of active learning also have been demonstrated for complex stimuli, such as associating a face with a name (Carpenter and DeLosh, 2005), as well as for learning foreign-language vocabulary (Pashler *et al.*, 2005). Active learning is especially effective when the correct answer is provided as feedback after an error (Bangert-Drowns *et al.*, 1991; Pashler *et al.*, 2005; Marsh *et al.*, 2012).

Interventions Used in the Program. The program described and tested here incorporates learning routines that were first developed by Gauthier (Gauthier and Tarr, 1997) for use in her studies of holistic visual processing. These routines use interleaving of identifications within a session, and have the potential to space repeated review over several days. They also incorporate active learning in the form of identification quizzes and tests. Quizzes provide immediate feedback on students' answers, while tests do not. Thus, the software effectively implements the SPRINT (spacing, retrieval, and interleaving) mechanisms studied by many cognitive psychologists in the past decade (McDaniel, 2012). The *Materials and Methods* section describes the software in more detail.

Hypothesis

On the basis of this background, we hypothesize that use of the software program Visual Learning—Plant Identification, which incorporates these features, will result in increased identification ability. We tested this hypothesis in a class on plant systematics during the Spring 2013 semester.

MATERIALS AND METHODS

Study Design

A within-subjects design was used (Table 1). Participants were 48 University of North Carolina at Greensboro (UNCG)

Table 1. Experimental design

Student group	Study set A	Study set B
1	VL-PI	Self-study
2	Self-study	VL-PI

students enrolled in two sections of a plant systematics course in Spring 2013. This is the first course in which these students were exposed to any significant botany component, and it is the only course at UNCG in which plant identification is taught. To our knowledge, none of the students had any prior experience in plant identification. An important learning goal of the course is for the students to be able to recognize common southeastern U.S. vascular plants. For this experiment, the 48 enrolled students were randomly assigned to two groups of twenty-four students each, equally divided between the two lab sections. Each group was assigned half of the required plants to learn with VL-PI and half to learn with other methods, described in *Additional Study Methods*. Of the 125 required genera and 69 required families, 97 genera and 51 families were included in VL-PI and added to one of the study sets. Study set A consisted of 48 genera in 28 families, while set B had 49 genera in 23 families (Supplemental Material 1). The discrepancy between sets was due to the fact that many families contain several genera, and it was decided not to split families between study sets. Each group of students received a different study set and was assigned different scripts (explained in the following section) to run as homework (Supplemental Material 2).

Description of the Program, VL-PI

VL-PI was designed to teach visual expertise in plant recognition (Curby and Gauthier, 2010). It is written in Java and requires installation of the Java Runtime environment. Although the program contains HTML-based tutorial and help files, students often benefit from instruction on how to use the program. This instruction was provided face-to-face in 2013, but we have used instructional videos to accomplish the same purpose in other semesters. While VL-PI focuses on plant identification, other similar programs have been developed on subjects as diverse as herpetology, plant life cycles, organic chemistry, and biochemistry (www.metisllc.com).

To be able to identify a plant, students must have a clear understanding of the taxonomic concept that underlies the species. This concept is usually acquired somewhat haphazardly through lecture presentations, exposure to the plants in the laboratory, and the use of keys (Walter and Winterton, 2007). VL-PI addresses this problem by exposing students to multiple images in an active-learning format. Exposure to variation has been shown to be important for concept learning (Fried and Holyoak, 1984; Medin, 1989; Perlman *et al.*, 2012). The software has a self-study mode, whose purpose is to familiarize students with the plants they will have to identify in the more active-learning modes of the program. Students are also able to select the taxonomic level at which they wish to learn the plants and whether they wish to learn them by scientific or common name. There are two preview mechanisms by which the students can become familiar with the taxa: they can preview images of the selected taxa via use of the arrow keys, with the taxon names superimposed (Figure 1A); or they can view them in a timed image display, either with their names superimposed or followed by their names.

The active-learning modes employ retrieval practice quizzes. Quiz design was adapted from the literature on holistic visual perception, which assures a tight linkage to the results presented in this literature (Gauthier and Tarr, 1997; Gauthier *et al.*, 1998; Tanaka *et al.*, 2005). Four types of quizzes are incorporated into the program. In the modes in which a user has to type a taxon name, spelling counts, but spelling fidelity is adjustable through an advanced program setting. In all modes, if the users get an answer correct, they are given positive reinforcement via a pop-up message that contains an affirmation of their success: "Great!," "Correct!," "Affirmative!," and similar. If they make an error, they are given the option of a redoing the question. The images used in VL-PI are standardized, so salient parts are easily visible and not obscured by extraneous features (Baskauf and Kirchoff, 2008). The four quiz types are:

1. Image naming with prompt: the user sees an image with the name of the image superimposed and responds by typing the name in a response box (Figure 1B).
2. Image naming without prompt: An image is displayed, the screen is cleared, and a response box appears (Figure 1C). the user types the name of the image at the selected taxonomic level into the response box. This type of quiz demands the most active participation from the student.
3. Image comparison: the user sees two images side by side (Figure 1D) and, after the screen clears, presses "y" if the images belong to the same taxon or "n" if they do not.
4. Image verification: The user sees an image (Figure 1E), then the screen is cleared, and, after a short delay, a possible identification appears (Figure 1F). The screen is cleared, a response box appears, and the user presses "y" if the name identifies the image and "n" if it does not.

The duration of the image display is adjustable. Permissible display times range from 0.1 to 4.0 s. Novice students begin with longer display times and progress to shorter times as their experience grows. Students can use VL-PI themselves, or instructors can create custom study/quiz sessions called scripts that use instructor-selected images and study/test routines. The scripts can be delivered to students by either by email or a content-delivery system such as Blackboard or Canvas. Output grading files are created when a script is run and can be returned to the instructor for inclusion in the student's course grade. A subsidiary, helper program allows the instructor to aggregate the grading files and easily give a grade to a set of scripts assigned as homework. Student compliance with learning tasks assigned outside a script is tracked in an output file keyed to the username of the student. This tracking file keeps a record of program use and user performance. The file tracks which routines are used, the duration of use, and, where applicable, the number of correct and incorrect responses. The results stored in this file can be displayed to the student, who can print them to submit to the instructor.

VL-PI is free and can be downloaded from B.K.K.'s company website, www.metisllc.com. A full description of the

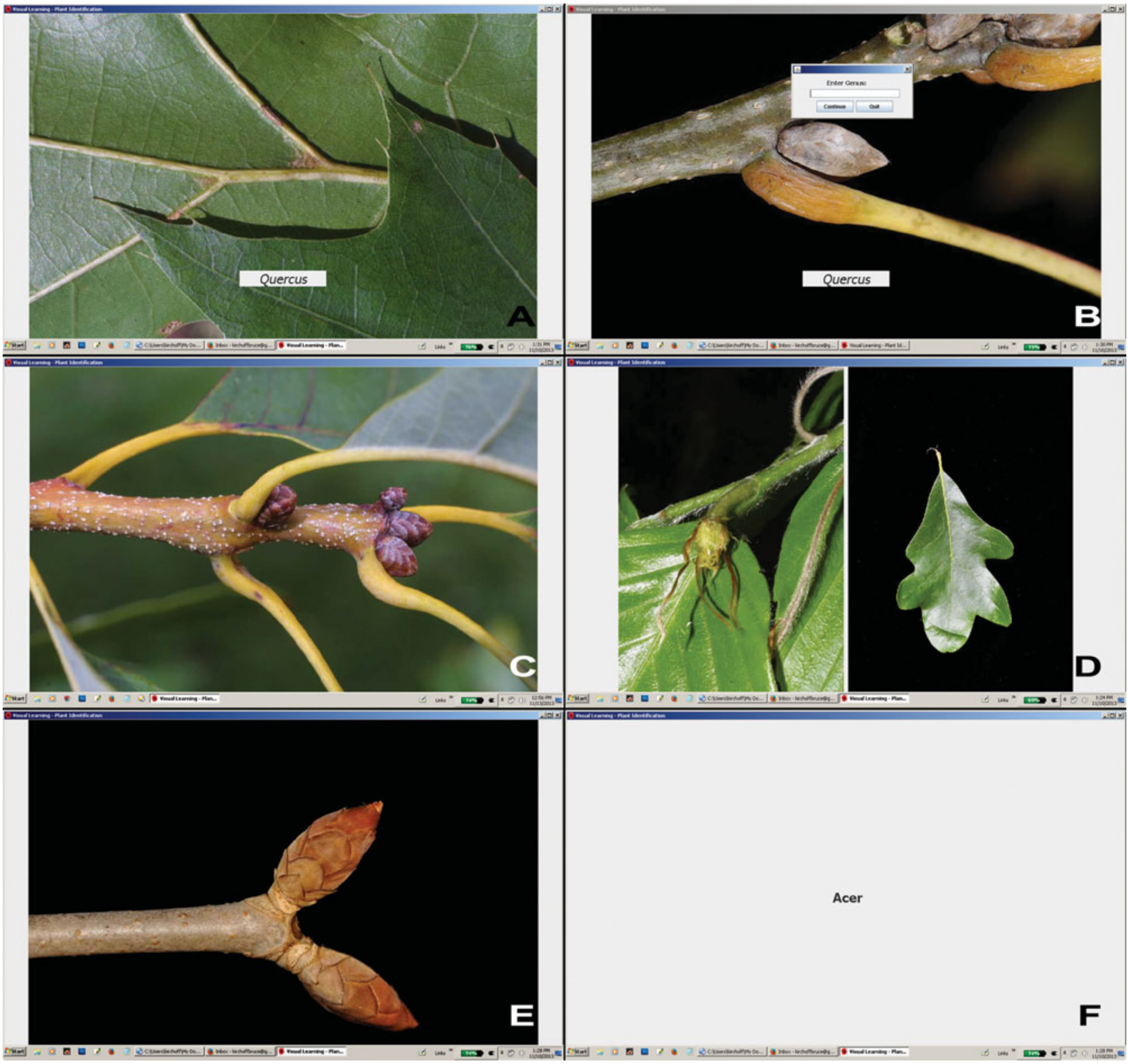

Figure 1. Screen shots of VL-PI. (A) Study taxa. The images are displayed with their names, and advanced with the arrow keys. (B) Image naming with prompt. This mode helps the student to associate the name of the taxon with its image by asking him or her to enter the taxon name in the response box while the image and its name remains on the screen. (C) Image naming without prompt. The image is displayed and then cleared from the screen before the response box appears. (D) Image comparison. Two images appear side by side and are then cleared from the screen. The response box appears, and the user enters "y" if the images are from the same taxon and "n" if they are not. (E and F) Image verification. An image appears (E) and is cleared and is then followed by one of the names (F) selected from the current set of taxa. The screen is cleared again, and the user responds "y" if the image and name match and "n" if they do not.

program and its features is available in the help and tutorial files, accessible from the Help menu in the program.

Experimental Scripts

VL-PI scripts were created for the two taxa sets: set A and set B. The complete list of scripts is given in Supplemental Material 2, broken down by week. Only genera were assigned in scripts. Scripts were not used to teach family identification, although these identifications were required and included in the final exam. When the same taxon was repeated across scripts, a new set of images was chosen for the student to learn. Some of the images overlapped with previous script sessions, and some were new.

Students completed the VL-PI assignments on their own time, as homework. Scripts were assigned through a content-delivery system (Blackboard), and the grading files were turned in through the same system. Homework, the majority of which was assigned with VL-PI, accounted for 15% of a student's course grade.

Additional Study Methods

The VL-PI scripts were added to an existing course in which students were encouraged to use active studying techniques, such as retrieval practice, outside the classroom. For the plants not studied with VL-PI, these were the only study methods available. For instance, the students were encouraged to find sets of diagnostic features (characters) for each taxon and to use them in creating flash cards and other study aids. The diagnostic features were presented first in lecture and then reinforced in the laboratory period. The students were told to begin solidifying their knowledge of the features in lab, where they could see and work with the living plants, and then to continue at home using their lab manual and textbook as study aids. The lab manual contains a list of the required taxa and their diagnostic features. Once they had established diagnostic feature lists, the students were given four methods to use to learn the taxa: 1) They were told to make a mental image of each taxon and to keep it in mind as they reviewed its features. Repeated work with this method taught the students to use the image as a visual cue to help them remember the distinguishing features. 2) The use of flash cards was recommended. Students were told to write the name of the taxon on one side, and its diagnostic features on the other. If some of these features were more important than others, students were told to highlight them so they could focus their study on these features. At least one student also made image-based flash cards with a picture of the taxon on one side and its name on the other. 3) To help associate diagnostic features with the taxa, students were encouraged to make feature-by-taxon lists. The first entry in the list is the feature, and the second entry is a list of the taxa that have this feature. For instance, if the feature were "microphylls," the list of taxa with this characteristic would be *Lycopodium*, *Selaginella*, and *Isoetes*. 4) Finally, it was suggested that students make taxon-by-feature tables. The row entries in these tables are the taxa, and the column entries are the diagnostic features. Although we observed students using all of these methods, it was not possible to know how many students used each technique, or how often they engaged in the recommended behaviors.

In addition to their textbook (Simpson, 2010), the students had the following resources available to them to assist their study: 1) all of the images from Simpson (2010), available on our content-management system as JPEG files; 2) a lab manual that listed the required taxa and gave their distinguishing characteristics; 3) access to the Internet and suggested websites that include images, taxonomic descriptions, and definitions of terms; 4) access to herbarium specimens and fresh and fluid-preserved specimens for a period of several weeks following each lab. The herbarium specimens and preserved materials were available throughout the semester, while the fresh specimens were only available for a week or two following the lab in which they were first used. Only a few students were observed making use of the herbarium specimens and fluid-preserved specimens.

The additional study materials that the students generated, if any, were not handed in or graded. They received no course credit for this work, above any that accrued through increased test performance.

Final Assessment

At the end of the semester, all of the students took the final exam. The exam included questions on the theoretical content of the course, and identification of plants from written descriptions, photographs, and living plants. The theoretical aspects were tested with a variety of questions including true–false, matching, short answer, and definitions of terms. The identification questions tested the students' ability to name the taxa from written descriptions, photographs projected with PowerPoint, and from living plants. Half of the identification questions were on plants from set A; the other half were from set B. There were also questions on plants that were not studied as part of either set. The identification questions from the final examination are available in Supplemental Material 3A.

The taxa selected for identification were chosen to avoid duplication in the three identification portions of the exam. Taxa were first selected for the living plant identifications based on what was available at that time of the year. Taxa were next selected for the photographic identifications, and finally for the written identifications. Approximately equal numbers of families and genera were included from both study sets (A and B) for each type of question. The complexity of the exam with equal numbers of taxa from two taxonomic levels (genus and family), two sets of taxa, selected for three types of questions, and taking time of year into consideration, allowed little leeway with respect to which plants were included in the exam.

The students had 3 h to complete the exam, which was divided into two parts. Part 1, which covered theory and written description and photographic questions (Supplemental Material 3A), lasted 2 h. Part 2, which covered live-plant identifications and use of an identification key, lasted 1 h (Supplemental Material 3A).

For the 30 photographic identifications, the photographs were projected with PowerPoint and a data projector. The images cycled continuously during the 2-h exam, with each slide remaining on the screen for 15 s. Each slide contained several images of the plant to be identified (Supplemental Material 3B). Students had to identify both the genus and the family, except in one question, in which they only had to identify the family. Of the 30 questions, there were 13 from set A, 13 from set B, and four from neither set. The family-only question was one of the questions from neither set. Images were obtained from the Internet through Google's image search. No images from VL-PI were used on the exam.

Twenty living plants were selected for identification on the final exam. These were all locally growing plants that the students could have seen around the campus before the exam. Students had to identify the genus and family of 17 of these plants, and only the family of three of them. Of the 20 plants, there were nine from set A, eight from set B, and three neither set. Of the three plants for which only a family name was required, one came from set A, one from set B, and one from neither set. Students could approach and touch the plants as often as needed during the exam period.

For the written description identifications, the students were asked to provide the correct taxon name based on a written description of the taxon. For example, for the genus *Vinca*, they were given the following description and asked to name the genus that fit this description: "Herbs with slender trailing stems; leaves opposite, simple, broadly lanceolate to ovate; flowers salverform, of five usually violet petals. Fruit a follicle, sometimes." Similar descriptions were used at the family level. The genus and family identifications were separated into different questions. In any one question students were asked to identify the genus or the family, but not both. Of the 14 written genus questions, six were from set A, six from set B, and two from neither set. Of the 10 family descriptions, four were from set A, four from set B, and two from neither set.

Grading

The identification questions on the final exam were graded by M.H. and a second lab instructor. The grading was verified by a student grader. Small spelling errors were allowed. A question would still be counted correct if one or two letters in the name were incorrect, as long as the taxon name was reasonably close to correct and the graders could easily interpret the intended identity of the plant. Common names were not accepted in lieu of the correct scientific name.

Student Assessment of VL-PI

Student judgments on the utility of VL-PI were captured with five, seven-point Likert-scaled questions administered anonymously on paper after the final exam. An answer of one (1) was coded as "strongly agree," and seven (7) was "strongly disagree." The questions were: 1) I thought that the program was easy to use. 2) I had to study less because I used this program. 3) I enjoyed using the program. 4) Using the program improved my grade. 5) The program made it easy to learn plant identification. A free-response box was also provided with the prompt, "Please tell us about your experience in using VL-PI." Student responses were collected in a sealed envelope and stored until after course grades were computed and turned in.

For assessment of the free responses, comments were first broken into idea units by B.K.K. The 75 idea units were then independently scored by B.K.K. and M.H. as positive comments, negative comments, neutral comments (such as suggestions for improving the program or factual statements about the program), and other (such as incorrect statements about the program; unclear statements; or comments about logistics, such as when the program was first made available to students). Interrater agreement was 80%. Discrepancies were resolved through discussion so that each idea unit received a single consensus score. For determining whether a student's overall response regarding the software was positive or negative, all neutral comments were thrown out, and a percent positive score was calculated for each student. Average scores below 1.5 were treated as positive; those above 1.5 were negative; and those of exactly 1.5 were scored as equivocal. The full set of written comments is available in Supplemental Material 4.

Retention of Learning

For assessment of retention, a retest was given 6 mo following the original exam. The retest was designed, administered, and graded by R.D.-J. Participants were approached about taking the retest in exchange for a $10 gift card. A total of 12 students participated in the retest: nine from group 1 and three from group 2. The test was administered on laptop computers using PowerPoint. Half of the taxa were from set A, and half were from set B (Supplemental Material 5). Only 12 of the taxa were reused from the original exam. New images, with which the students had no prior experience were used for all questions. As with the final exam, each test slide asked for the family and genus, and presented up to three color photographs of the taxon. Students wrote their answers on a separate answer sheet.

Before taking the test, students were allowed 15 min to review a list of all required family and generic names from the course, arranged alphabetically by genus. No images were provided for review. Students were allowed 1 h to complete the retest. Due to scheduling conflicts, two students were emailed the exam materials and allowed to take the test on their home computers.

Statistical Analyses

All statistical analyses were carried out with SPSS version 19 or 21. For the final exam results, a paired-samples t test was used to check for differences in means between exam performance on the taxa that were studied with VL-PI and those that were not. Each question type (photographs, living plants, written descriptions) was tested separately at the genus and family levels. For the delayed retest, a 2 Test Occasion (Final versus Delayed) × 2 Item Type (VL-PI vs. not in VL-PI) repeated-measures analysis of variance was used. Follow-up tests were conducted using paired-samples t tests.

RESULTS

For each identification type, we compared the student's percent accuracy on the taxa that were learned with VL-PI with their accuracy on those that were not (Table 2). Students showed a statistically significant 25% advantage (6.25 questions; $p < 0.001$) in identifying the genus from the photographs (Table 2). Interestingly, they also showed a 17% statistically significant advantage (4.25 questions; $p < 0.001$) when identifying the family from photographs, although families were never explicitly the object of training. To verify this result, we removed all questions in which the genus and family names were based on the same roots (e.g., *Psilotum* and Psilotaceae, *Ginkgo* and Ginkoaceae, etc.) and reran the analyses (Table 3). The results showed a similar (15%; 2.1 questions) and statistically significant ($p < 0.001$) advantage for families from photographs.

Identification of the genus of living plants was also statistically significantly improved (8%; 1.36 questions; $p = 0.007$) through use of the program (Table 2). Identification of the family for living plants was improved numerically (5%) for both the full (0.85 questions) and reduced data sets (0.55 questions; Tables 2 and 3), but neither result was statistically significant.

Identification of the genus from written descriptions was statistically significantly superior (13%; 1.56 questions; $p = 0.002$) after training (Table 2). There was no significant improvement for family identifications from written descriptions, although numerically the advantage (6%; 0.48

Table 2. Final exam, mean percent correct for taxa practiced with VL-PI versus those learned with other methods ($n = 46$)

	VL-PI (% ± SE)	Other (% ± SE)	VL-PI − Other (%)	Significance
Genus				
Photographs	80 ± 2.4	55 ± 2.8	25	$p < 0.001$
Living plants	81 ± 2.9	73 ± 3.4	8	$p = 0.007$
Written text	67 ± 4.1	54 ± 4.3	13	$p = 0.002$
Family				
Photographs	71 ± 2.7	54 ± 3.1	17	$p < 0.001$
Living plants	77 ± 3.5	72 ± 3.7	5	$p = 0.093$
Written text	44 ± 4.4	38 ± 4.7	6	$p = 0.209$

questions; p = 0.209) suggested an improvement due to the use of VL-PI (Table 2). Because the genus and family written questions were separated on the exam (Supplemental Material 3A), there can be no question of similarity in names influencing performance on the written questions.

Retention of Learning

For the genera (Figure 2A), there was a significant main effect of time ($p < 0.001$), indicating that the students performed overall more poorly on the retest than the final exam (35% vs. 67%; 6.4 questions). There was also a main effect of using VL-PI ($p < 0.001$; 64% vs. 38%; 5.2 questions), and a significant interaction ($p = 0.024$). Follow-up t tests indicated a significant advantage of VL-PI on both the final exam ($p < 0.001$)—replicating the earlier results with the subset (Table 3)—and on the delayed retest ($p = 0.007$).

For families, the pattern was similar (Figure 2B). There was a significant main effect of time ($p < 0.001$), indicating that the delayed retest accuracy was lower than the initial test (29% vs. 60%; 6.2 questions). Families in VL-PI were better identified than families that were not ($p < 0.001$; 55% vs. 35%; 4 questions). The interaction at the family level was not statistically significant ($p = 0.251$). The effect of VL-PI remained significant when a t test was run only on the delayed retest data ($p = 0.012$).

Student Evaluations of the Software

The postexam survey showed that students were generally positive about using the software. The Likert-scored means were all in the positive range (Table 4). The mean for question 5, "The program made it easy to learn plant identification," was 2.51 on a scale where 1 was "completely agree" and 7 was "completely disagree," indicating a general satisfaction with the program. For the free responses, there were 75 statements by 37 students. One student's comment was judged neutral and was excluded from the following percentages. The percentage of remaining students offering favorable opinions was 75% (27 students), with 14% negative (5 students), and 11% equivocal (equal numbers of positive and negative comments; 4 students).

The comments (Supplemental Material 4) suggest that the students thought that the use of program enhanced their learning. An analysis of the few students judged to have made negative comments illustrates this point: Student 13 stated "Needs to be more user friendly. Where you don't have to upload scripts" and rated the program 3 on question 5, indicating that he or she was more satisfied than not with the program. This negative comment deals with logistics and not with learning. Student 18 stated "Only complaint: When 0.3 s show-time, it felt as if I was memorizing a background, not identifying a plant. I felt longer times were more beneficial to learning characteristics" but again gave the program a 3 on question 5. Student 22 stated "Suggestions: - Text that gives 'affirmative' 'you got it'—change to reiterating family & genus etc.—tested on genus-had difficulty connecting visually to family" but gave it a 2 on question 5, indicating above average satisfaction with the program. Student 28 stated "I think sometimes I memorized a picture rather than the plant and its characteristics" and also rated the program a 2 on question 5. Finally, student 40 noted "It was not helpful, and very time consuming," which is distinctly negative, but rated the program 4, in the middle of the range, on question 5. Of the three students who gave the program the worst rating (7) on question 5, two (Students 10 and 23) made no comments. The third, Student 42, stated "It helped a lot w/memorizing different plants."

Table 3. Reanalysis of data including only exam scores for questions where genus and family names were different ($n = 46$)

	VL-PI (% ± SE)	Other (% ± SE)	VL-PI − Other (%)	Significance
Genus				
Photographs	83 ± 3.2	52 ± 4.3	31	$p < 0.001$
Living plants	82 ± 2.5	70 ± 3.3	12	$p = 0.001$
Family				
Photographs	66 ± 4.2	51 ± 4.6	15	$p < 0.001$
Living plants	72 ± 3.4	67 ± 3.7	5	$p = 0.173$

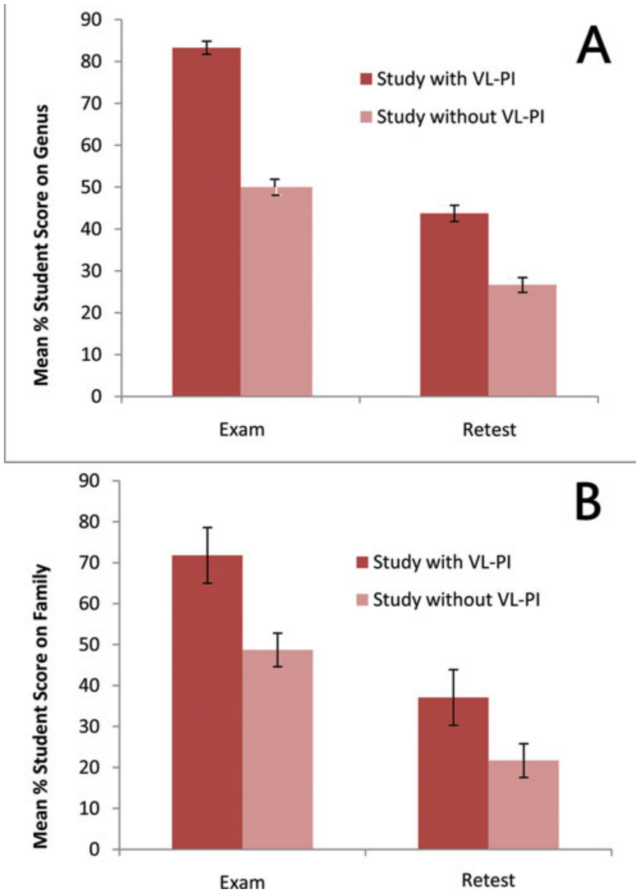

Figure 2. Percent correct identifications on the final exam for the course and the delayed retest. Error bars represent ± SE. (A) Responses at the genus level ($n = 12$). (B) Responses at the family level ($n = 12$).

DISCUSSION

The Importance of Plant Identification

Plant identification is an important skill in the fields of natural resource management, agriculture, and horticulture (Convention on Biological Diversity, 1996; Stevens, 2001; Terlizzi *et al.*, 2003; Arteca, 2006; Mangold and Parkinson, 2013). Correct identification affects every aspect of how plants are conserved and used in the landscape. Rare and threatened plants can be conserved only if they are recognized as such, which requires a comprehensive knowledge of the flora on at least a regional basis. Bioprospecting and the sustainable use of biodiversity depend first on accurate identifications (Convention on Biological Diversity, 1996). Without reliable identification it is impossible to repeat collections of the same species in order to confirm the existence of promising compounds. No amount of chemical analysis can produce effective resource utilization if the identification is incorrect. Plant identification is important for weed control in agriculture and for effective natural resource utilization in range management. Effective weed control is based on the properties of the specific weed. Some respond well to chemical treatment, while others require manual removal. Rangeland wildlife management is based on having the correct mix of plants to provide food and shelter for wildlife. Plant identification plays an essential role in this process. In horticulture, time of planting, time and method of pruning, and fertilization schedule all depend on the identity of the plant. Because horticultural plants are not always distributed under the correct name, plant identification is a critical skill for every student of horticulture.

Plant identification is usually taught at the university level, either as part of a plant systematics course or in a course in field botany. Horticulture and agriculture courses also have important plant identification components, as do animal and veterinary science courses that emphasize poisonous plants (Burrows *et al.*, 2014). These courses use a mixture of the techniques described in the following paragraphs for systematics and field botany courses. Although field botany used to be a staple of botany departments, the merger of botany and zoology into biology and the increasing emphasis on quantitative skills (National Research Council, 2003) have led to a reduction in institutions offering field identification courses. What plant identification is taught usually occurs as part of a course on plant systematics. Most of these courses place an equal emphasis on learning plant identification at the family level, and on conceptual issues related to the history and practice of systematics. Theories of classification, phylogeny reconstruction, basic plant morphology, and plant reproduction are often covered in addition to plant identification (Judd *et al.*, 2008; Simpson, 2010).

Within systematics courses, there are two ways by which plant identification is commonly taught. The first method involves focusing on diagnostic characters. The instructor teaches students to search for and recognize suites of features that are characteristic for each taxon and that distinguish it from other similar taxa. Once the student has developed a good memory of the features of the taxa (usually families), he or she is taught to evaluate unknowns in the following way. "I can tell that this plant is a member of the family Aquifoliaceae because it is a woody plant with alternate leaves, minute triangular, black stipules, its flowers are imperfect with the flat stigma arising directly from a globose ovary, and it has

Table 4. Results of Likert-scored questions (1 = strongly agree; 7 = strongly disagree)

	1	2	3	4	5	6	7	Mean	SD
I thought that the program was easy to use	17	14	5	3	0	2	1	2.17	1.48
I had to study less because I used this program	4	8	14	5	4	1	7	3.65	1.88
I enjoyed using the program	7	16	8	6	3	2	1	2.81	1.50
Using the program improved my grade	17	9	6	4	3	1	2	2.48	1.73
The program made it easy to learn plant identification	12	14	9	5	0	0	3	2.51	1.58

colorful drupes with several pits." This method depends on a clear understanding and memory of the technical terms.

The second method consists of teaching the students to use a taxonomic key and giving them extensive practice in identifying unknown plants with the key. Keys are instructions for identifying unknown taxa. The user makes a series of choices about the unknown organism's characteristics and, if the choices are correct, is led through a series of steps to the organism's identity. In using keys, students are forced to be observant and must learn the technical terms upon which the keys depend. After keying out several members of the same family, a student, in theory, begins to see family-level resemblances and eventually develops a mental concept of the family. He or she begins to see suites of common characters and builds up knowledge of plant diversity at the family level.

Both of these approaches depend on a student's ability to learn the technical and often arcane terminology of plant identification. The learning curve for these terms is very steep. A partial list of the terms for leaf shape will demonstrate the difficulty: linear, oblong lanceolate, elliptic, oblanceolate, ovate, broadly elliptic, obovate, orbicular, reniform. There are separate sets of terms for the shape of the leaf apex, the leaf base, and the margin; the texture of the leaf and its venation and covering of hairs; not to mention the terms associated with the other parts of the plant. Needless to say, only students who see a direct application for this knowledge in their careers find the motivation to do well in courses that depend on a detailed knowledge of this terminology.

Field botany courses use a very different approach. The emphasis in these courses is on field trips and identification of the plants in their natural habitats. The instructor points out and identifies the plants each time they are seen and usually mentions a few of the distinguishing characteristics that help identify the plant. While some technical terms may be used, the emphasis is on learning to recognize species by sight. The emphasis is on the species, not the genus or family levels. Cut specimens are often brought into the lab for review, and a list of plants seen on each field trip is often provided to the students for later study.

In all of these approaches, most of the learning must take place outside the classroom. The techniques used by the control group in this study are typical of the techniques used for this purpose.

The Advantage of Using VL-PI

Guided use of VL-PI through a series of structured assignments (scripts) statistically significantly enhanced students' ability to identify plants. The benefits of using VL-PI were sustained even after a 6-mo delay. The effects were largest for information similar to the training materials—the genus identified from photographs—but there was also a benefit in identifying the family from photographs, which was never explicitly trained. Most likely these gains were mediated by better identification of genera, which enabled the students to remember the family. Furthermore, there were statistically significant gains in identifying genera of living plants and genera from written descriptions. Neither of these skills was directly trained with the software.

Although the percent benefit was smaller for living plants than for photographs, it is instructive to note that the baseline scores on living plants were higher overall than for photographs (Table 2). Students scored better than 70% (12 out of 17) correct on the living plants that were not studied with VL-PI, and so there was limited room for improvement after the training. The higher baseline could be due to several possible reasons. First, the living plants are all from the region in which the students live, suggesting that those plants are likely to be more familiar to them than the more disparate examples chosen for photographic evaluation. Second, the living plants may have been easier to identify, because students could examine them at their leisure, handle the plants, and see more of the plant than just what was displayed in the photographs. There may have been clues that could not be discerned from the photographs that helped them identify the plants.

The results presented here lead us to conclude that VL-PI offers a validated solution to problems in plant systematics education. The use of a within-subjects experimental design eliminates many types of possible influences on the outcome. For instance, it is theoretically possible that the students had prior knowledge of plant identification and that this knowledge influenced the results we obtained. However, for this to be true, students in group 1 would have had to have been familiar with the plants in set A, but not those in set B, before they entered the course. Students in group 2 would have had to have prior knowledge of the plants in set B, but not those in set A. This is so unlikely that we feel the possibility of prior knowledge affecting the outcome can be disregarded. The same is true of many other types of effects that can influence the outcome of between-subject designs.

One factor that could affect the outcome is time on task. It is possible that students spent more time studying the plants assigned with VL-PI then those in the other study set. In fact, VL-PI was designed especially for the purpose of increasing study time in plant systematics courses, so it is entirely possible that the students did study more because of the assigned scripts. However, to attribute the whole of the effects to increased study time would mean ignoring the large psychological literature that establishes the efficacy of these techniques. VL-PI takes advantage of several well-established learning principles, including development of perceptual expertise, interleaving of examples to train category identification, spacing of practice, and testing effects (see *Introduction*). In future work, it would be helpful to compare VL-PI with alternative study methods that are carefully monitored, as this was not possible in the current study. It would also be interesting to examine the effectiveness of VL-PI with more experienced students who are deepening their knowledge of plant identification.

Although creating weekly scripts can take up to 2 h per week, assignments can be given without scripts and student compliance monitored through the use of printed screen shots of the progress screen available in the program. When used in this way, assignments require only minimal time to prepare and grade. Use of VL-PI requires only about three quarters of an hour a week of instructor time to implement if assignments are given without scripts, and about 1.5–2 h of time if scripts are used.

To our knowledge, this is the first study to show transfer of learning from lower- to higher-order categorization levels (genus to family). From a cognitive psychology perspective, this is a worthwhile empirical finding on its own. One method by which this could work is if students have memorized the hierarchical relationships between genera and families, and therefore, when they can identify the genus, they are often able to identify the family as well. For some plants, the naming of the family is related to the naming of the genus, because the names share the same roots, but the removal of the data for taxa in which the genus and family name are the same shows that the benefits are not limited to these cases. This suggests that learning to categorize at the lowest level provides a benefit to categorization at higher levels, at least when the stimuli are difficult to learn.

A barrier to incorporating principles from cognitive psychology into the college classroom is that natural scientists rarely have the expertise to analyze the learning outcomes they want to achieve and to design appropriate assignments. Similarly, cognitive psychologists rarely have the disciplinary knowledge necessary to identify the most critical learning goals in the natural sciences. However, collaborative teams can produce effective learning tools that can significantly enhance learning outcomes. These teams are also well suited to diagnose why a particular learning method is or is not working.

Although this study explores perceptual learning in plant identification, there is no reason that programs like VL-PI cannot be created to support different category-learning domains. A sister program has recently been produced to teach herpetology, and two related programs exist in areas of chemistry: organic chemistry functional groups and amino acid structure (www.metisllc.com). Extensions to other fields merely require new images and image categorization to be inserted into the program framework. There is no reason to think that programs created for other disciplines will not produce learning outcomes similar to those achieved with VL-PI. Significant learning gains should be possible in any discipline that depends on visual data. However, it is important to be aware that there is at least one exception to the learning advantages gained from interleaving members of different categories. When only a few features discriminate between otherwise highly similar categories, presenting the images of a single category in a block has been shown to be more effective than interleaving examples. In these cases, the problem is one of identifying the few shared features that distinguish a category. This is better accomplished through repeated exposure to the distinguishing features of the category, displayed in blocks of images (Goldstone, 1996; Carvalho and Goldstone, 2014). It remains to be seen whether any natural science domains will have this characteristic.

ACKNOWLEDGMENTS

The authors thank Steve Baskauf (http://bioimages.vanderbilt.edu), Martin Wall, Nadia Talent, and Rodger Evans for providing images for use in VL-PI. We also thank Isabel Gauthier for her help in understanding the visual-processing literature during the early stages of work on this project. The first version of the program was published by Missouri Botanical Garden Press under the title *Woody Plants of the Southeastern United States: A Field Botany Course on CD* (www.mbgpress.info/index.php?task=id&id=08025). We thank Victoria Hollowell, the scientific editor and head of the press, for her support of the project. Partial funding for the development of VL-PI was obtained from two Advances in Teaching and Learning grants from the Teaching and Learning Center (now the Faculty Teaching and Learning Commons) of UNCG. Additional funding was obtained from Charles Sturt University, Wagga Wagga, Australia, through the auspices of Dr. Geoffrey Burrows. This research (study 09-0163) was reviewed by the UNCG institutional review board and was determined to be exempt from further review under the Code of Federal Regulations 45 CFR 46.101(b).

REFERENCES

Arteca RN (2006). Introduction to Horticulture Science, Florence, KY: Thomson Delmar Learning.

Bahrick HP (1979). Maintenance of knowledge—questions about memory we forgot to ask. J Exp Psychol Gen *108*, 296–308.

Bahrick HP, Bahrick LE, Bahrick AS, Bahrick PE (1993). Maintenance of foreign-language vocabulary and the spacing effect. Psychol Sci *4*, 316–321.

Bahrick HP, Phelps E (1987). Retention of Spanish vocabulary over eight years. J Exp Psychol Learn *13*, 344–349.

Bangert-Drowns RL, Kulik C-LC, Kulik JA, Morgan M (1991). The instructional effect of feedback in test-like events. Rev Educ Res *61*, 213–238.

Baskauf SJ, Kirchoff BK (2008). Digital plant images as specimens: standards for photographing living plants. Vulpia *7*, 16–30. www.cals.ncsu.edu/plantbiology/ncsc/vulpia/pdf/Baskauf_&_Kirchoff_Digital_Plant_Images.pdf (accessed 9 June 2014).

Bukach CM, Gauthier I, Tarr M (2006). Beyond faces and modularity: the power of an expertise framework. Trends Cogn Sci *10*, 159–166.

Burrows GE, Krebs GL, Kirchoff BK (2014). "Visual Learning—Agricultural Plants of the Riverina"—a new application for helping veterinary students recognise poisonous plants. Biosci Educ, DOI: 10.11120/BEEJ.2014.00028.

Carpenter SK, Cepeda NJ, Rohrer D, Kang SHK, Pashler H (2012). Using spacing to enhance diverse forms of learning: review of recent research and implications for instruction. Educ Psychol Rev *24*, 369–378.

Carpenter SK, DeLosh EL (2005). Application of the testing and spacing effects to name learning. Appl Cogn Psychol *19*, 619–636.

Carvalho PF, Goldstone RL (2014). Putting category learning in order: category structure and temporal arrangement affect the benefit of interleaved over blocked study. Mem Cogn *42*, 481–495.

Convention on Biological Diversity (1996). Practical Approaches for Capacity Building for Taxonomy, Report No. UNEP/CBD/SBSTTA/2/5, New York: UN Environment Program. www.iisd.ca/biodiv/sbstta/sb205.pdf (accessed 9 June 2014).

Curby KM, Gauthier I (2010). To the trained eye: perceptual expertise alters visual processing. Topics Cogn Sci *2*, 189–201.

Delaney PF, Verkoeijen PPJL, Spirgel A (2010). Spacing and testing effects: a deeply critical, lengthy, and at times discursive review of the literature. Psychol Learn *53*, 63–147.

Dempster FN (1987). Effects of variable encoding and spaced presentations on vocabulary learning. J Educ Psychol *79*, 162–170.

Farah MJ, Wilson KD, Drain M, Tanaka JN (1998). What is "special" about face perception? Psychol Rev *105*, 482–498.

Fermanian TW, Barkworth M, Liu H (1989). Trained and untrained individuals ability to identify morphological characters of immature grasses. Agronomy J *81*, 918–922.

Fried LS, Holyoak KJ (1984). Induction of category distributions—a framework for classification learning. J Exp Psychol Learn *10*, 234–257.

Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000). Expertise for cars and birds recruits brain areas involved in face recognition. Nat Neurosci *3*, 191–197.

Gauthier I, Tarr MJ (1997). Becoming a ''greeble'' expert: exploring mechanisms for face recognition. Vision Res *37*, 1673–1682.

Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (1999). Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. Nat Neurosci *2*, 568–573.

Gauthier I, Williams P, Tarr MJ, Tanaka J (1998). Training "greeble" experts: a framework for studying expert object recognition processes. Vision Res *38*, 2401–2428.

Goldstone RL (1996). Isolated and interrelated concepts. Mem Cogn *24*, 608–628.

Hole GJ (1994). Configurational factors in the perception of unfamiliar faces. Perception *23*, 65–74.

Judd WS, Campbell CS, Kellog EA, Stevens PF, Donoghue MJ (2008). Plant Systematics: A Phylogenetic Approach, 3rd ed., Sunderland, MA: Sinauer.

Kang SHK, Pashler H (2012). Learning painting styles: spacing is advantageous when it promotes discriminative contrast. Appl Cogn Psychol *26*, 97–103.

Karpicke JD, Blunt JR (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. Science *331*, 772–775.

Kornell N, Bjork RA (2008). Learning concepts and categories: is spacing the "enemy of induction"? Psychol Sci *19*, 585–592.

Kornell N, Castel AD, Eich TS, Bjork RA (2010). Spacing as the friend of both memory and induction in young and older adults. Psychol Aging *25*, 498–503.

Mangold J, Parkinson H (2013). Plant Identification Basics, Pamphlet no. MT201304AG, Bozeman: Montana State University Extension. www.msuextension.org/invasiveplantsMangold/documents/Publications_bulletins/mt201304AG_plant%20ID.pdf (accessed 9 June 2014).

Marsh EJ, Lozito JP, Umanath S, Bjork EL, Bjork RA (2012). Using verification feedback to correct errors made on a multiple-choice test. Memory *20*, 645–653.

McDaniel MA (2012). Put the SPRINT in knowledge training: training with spacing, retrieval, and interleaving. In: Training Cognition: Optimizing Efficiency, Durability, and Generalizability, ed. AF Healy and LEJ Bourne, New York: Psychology Press.

McDaniel MA, Agarwal PK, Huelser BJ, McDermott KB, Roediger HL (2011). Test-enhanced learning in a middle school science classroom: the effects of quiz frequency and placement. J Educ Psychol *103*, 399–414.

Medin DL (1989). Concepts and conceptual structure. Am Psychol *44*, 1469–1481.

National Research Council (2003). BIO2010: Transforming Undergraduate Education for Future Research Biologists, Washington, DC: National Academies Press.

Pashler H, Cepeda NJ, Wixted JT, Rohrer D (2005). When does feedback facilitate learning of words? J Exp Psychol Learn *31*, 3–8.

Pashler H, Rohrer D, Cepeda NJ, Carpenter SK (2007). Enhancing learning and retarding forgetting: choices and consequences. Psychon Bull Rev *14*, 187–193.

Perlman A, Hahn U, Edwards DJ, Pothos EM (2012). Further attempts to clarify the importance of category variability for categorisation. J Cogn Psychol *24*, 203–220.

Roediger HL, Karpicke JD (2006). Test-enhanced learning: taking memory tests improves long-term retention. Psychol Sci *17*, 249–255.

Schussler EE, Link-Prez MA, Weber KM, Dollo VH (2010). Exploring plant and animal content in elementary science textbooks. J Biol Educ *44*, 123–128.

Schussler EE, Olzak LA (2008). It's not easy being green: student recall of plant and animal images. J Biol Educ *42*, 112–118.

Simpson MG (2010). Plant Systematics, San Diego, CA: Elsevier.

Sobel HS, Cepeda NJ, Kapler IV (2011). Spacing effects in real-world classroom vocabulary learning. Appl Cogn Psychol *25*, 763–767.

Stagg BC, Donkin M (2013). Teaching botanical identification to adults: experiences of the UK participatory science project "open air laboratories." J Biol Educ *47*, 104–110.

Stevens R (2001). Plant Identification: Is It Worth the Effort? Ardmore, OK: Samuel Roberts Noble Foundation. www.noble.org/ag/wildlife/plantidworthit (accessed 6 March 2014).

Tanaka JW, Curran T, Sheinberg DL (2005). The training and transfer of real-world perceptual expertise. Psychol Sci *16*, 145–151.

Terlizzi A, Bevilacqua S, Fraschetti S, Boero F (2003). Taxonomic sufficiency and the increasing insufficiency of taxonomic expertise. Mar Pollut Bull *46*, 556–561.

Tilling S (1984). Keys to biological identification—their role and construction. J Biol Educ *18*, 293–304.

Vlach HA, Sandhofer CM, Kornell N (2008). The spacing effect in children's memory and category induction. Cognition *109*, 163–167.

Wahlheim CN, Dunlosky J, Jacoby LL (2011). Spacing enhances the learning of natural concepts: an investigation of mechanisms, metacognition, and aging. Mem Cogn *39*, 750–763.

Walter DE, Winterton S (2007). Keys and the crisis in taxonomy: extinction or reinvention? Annu Rev Entomol *52*, 193–208.

Wandersee JH, Schussler EE (1999). Preventing plant blindness. Am Biol Teach *61*, 82–86.

Wisniewski EJ (2002). Concepts and categorization. In: Stevens' Handbook of Experimental Psychology, vol 2., Memory and Cognitive Processes, ed. D Medin, New York: Wiley, 467–531.

Young AW, Hellawell D, Hay DC (1987). Configurational information in face perception. Perception *16*, 747–759.