Adv Physiol Educ 37: 233–241, 2013;
doi:10.1152/advan.00104.2012.

Process-oriented guided-inquiry learning improves long-term retention of information

Thea Vanags,[1] Kristen Pammer,[2] and Jay Brinker[2]

[1]*Office of Research, University of The Sunshine Coast, Maroochydore, Queensland, Australia; and* [2]*The Australian National University, Canberra, Australia*

Submitted 20 July 2012; accepted in final form 8 July 2013

Vanags T, Pammer K, Brinker J. Process-oriented guided-inquiry learning improves long-term retention of information. *Adv Physiol Educ* 37: 233–241, 2013; doi:10.1152/advan.00104.2012.—Many chemistry educators have adopted the process-oriented guided-inquiry learning (POGIL) pedagogy. However, it is not clear which aspects of POGIL are the most important in terms of actual learning. We compared 354 first-year undergraduate psychology students' learning in physiological psychology using four teaching methods: control, POGIL, POGIL without reporting [no report out (NRO)], and POGIL run by untrained graduate students [new facilitator (NF)]. Student activities were identical across POGIL variations and highly similar for control. Participants' knowledge was evaluated before (pretest), immediately after (posttest), and 2 wk later (followup). Control and POGIL groups showed no improvement at posttest, whereas NRO and NF groups both recalled more material than at pretest ($P = 0.002$ and $P < 0.0005$, respectively). In a surprise test 2 wk later, control ($P < 0.0005$), NRO ($P = 0.03$), and NF ($P < 0.0005$) groups recalled less than at posttest. The POGIL group showed the smallest drop in knowledge ($P = 0.05$). Importantly, the control group's knowledge was below pretest levels ($P < 0.0005$), whereas the POGIL, NRO, and NF groups' knowledge was not. Self-assessment of knowledge was consistent across groups at pretest, but POGIL participants had the lowest confidence at posttest and 2 wk later. At followup, the control, NRO, and NF groups showed greater confidence in their knowledge than the POGIL group ($P = 0.03$, $P = 0.002$, and $P = 0.004$, respectively). POGIL and its variations appear to consolidate existing knowledge against memory decay even when student confidence does not match performance.

process-oriented guided-inquiry learning; guided inquiry learning; long-term retention

WHEN THE NATIONAL SCIENCE FOUNDATION proposed a shift from faculty-centered to student-centered teaching in 1994–1995 (18), many chemistry educators in the United States took up the challenge to develop new methodologies to replace the age-old method of teaching by telling (5, 8). Many educators have embraced process-oriented guided-inquiry learning (POGIL), and its effectiveness has been evaluated extensively against traditional teaching methods (5, 9, 10, 12, 17, 21). The majority of these studies have shown improved learning outcomes for students as a result of POGIL (5). The POGIL methodology can be implemented by replacing some or all lectures and/or recitation sessions with POGIL workshops (9, 17). Although most evaluations of POGIL have been in the discipline of chemistry, there have been a small number of studies in other disciplines, such as high school biology and tertiary undergraduate biology, physics, mathematics, computer science, engineering, environmental science, marketing, education, and anatomy and physiology (2, 5, 8).

Knowing which components of the methodology contribute most to student learning is crucial because POGIL involves considerable effort on the part of both the facilitator and the student (2, 12). Facilitators must expend time and effort developing activities that mirror the learning cycle: exploration, concept formation, and application. This involves changing the way we present information, developing models for student inquiry and writing questions that build on students' prior knowledge and then lead students to make inferences in the construction of their own mental representations of principles and concepts (2, 10, 11). For students, the change from passive to active learning is usually enjoyable, but it requires greater effort by students in the classroom and the change from a didactic approach they are used to is not always welcomed by students, particularly if the benefits of the new approach are unclear (2, 5).

One of the key differences from traditional didactic teaching is the use of student teams in learning. Students often have limited (if any) experience in learning teams. Without assigned roles or specific direction within teams, students who realize they can do the task easily are likely to do it quickly, whereas those who cannot do the task will struggle on their own (10). Simply placing students in groups with the instruction to work together on a problem does not produce the benefits of team learning instrumental to cooperative learning (10). Such benefits can be achieved by the facilitator diversifying teams by assigning students to teams of three or four students rather than letting students self-select teams (9, 12) and by adopting POGIL-specific team roles to ensure positive interdependence (for the team to succeed, each person must succeed) of the team (10). Each of these roles needs to be critical to the team's success, and it is important to provide students with clear definitions of those roles (10).

The instructor's role as facilitator also differs from the traditional instructor role of providing information and answers to students' questions with definitive statements about content (7, 17). In POGIL, the facilitator's role is to support the students' construction of knowledge rather than providing the answers (9). If the facilitator regularly provides answers, students learn to wait passively for information instead of constructing concepts and principles themselves (10). Discoveries that students make themselves are better retained and lead to improved conceptual understanding because students think critically about the material and construct an understanding that addresses gaps specific to their mental representations (10). In POGIL sessions, the facilitator moves among students listening and watching them work (9). This provides the facilitator with a good sense of how students are performing

Address for reprint requests and other correspondence: T. Vanags, Univ. of The Sunshine Coast, Rm. B1.07A, Maroochydore 4556, Queensland, Australia (e-mail: tvanags@usc.edu.au).

(10) and allows the facilitator to judge whether an intervention in the form of a short, just-in-time lecture to clarify concepts is necessary (5).

As with any teaching technique, facilitators improve with experience. Faculty members who adopt an alternative pedagogy such as POGIL will most likely have reviewed the literature or attended a workshop before making this change. It is often attendance at such workshops that lead faculty members to consider making this change (5). But, when POGIL is used to replace recitation sessions, the facilitator may be a graduate student with little or no training in teaching and learning, no time to review the literature, and limited experience managing classes and learning teams. Evaluations of POGIL to date have not included implementation by graduate students meeting these criteria. In the published evaluations, POGIL sessions facilitated by graduate students involved those who were trained in the methodology. This training was (appropriately) quite extensive. For example, in the Lewis and Lewis study (14), peer leaders (other undergraduate students) were given training for 2 h/wk with experienced lecturers; the training involved working through the POGIL activities for the next session, which led to better performance by these students than those taught in traditional lectures. Hanson (9), in keeping with his argument that assistants need to be trained both in POGIL and learning theory and pedagogical approaches, ensured that graduate students at Stony Brook University received extensive training in pedagogy. Chemistry graduate students did three 3-h sessions the week before the semester started and attended weekly 2-h meetings during the semester. These sessions covered the subject material, administrative teaching matters, and pedagogy (9), a model of graduate student training that many universities would do well to emulate. Unfortunately, funding issues limit the resources for training, and many graduate student facilitators of recitation sessions have no training in pedagogical issues unless they seek out this training themselves (11). It is not known whether methodologies such as POGIL are undermined by a lack of training of teaching assistants (13) and to what extent that lack of training affects student learning outcomes, particularly for facilitation, as opposed to traditional teaching.

Finally, another core component of a POGIL session is that the teams' presenters are asked to "report out" their teams' answers on a whiteboard or a blackboard (9, 17). This can be done efficiently with presenters from each group simultaneously writing answers to different questions on the board (9). In keeping with the concept of students constructing their own knowledge, the facilitator then asks the class for agreement or disagreement on each answer (9). Reporting out could be considered a critical component of POGIL because it is through the resolution of differences between groups' answers that students engage in deeper learning and question their understanding of the material while explaining and arguing their case to their peers (1, 10). Although students will have engaged in some form of explanation and discussion within their learning teams, doing this across groups in the report out activity requires teams to justify their inferences and constructed knowledge at a higher level, thereby clarifying concepts (9). Higher-order learning encourages students to engage with the material at a deeper, more integrated level, leading to better long-term retention (1).

While important factors in POGIL implementation have surfaced in the literature, the importance of other factors in producing improved learning outcomes for students remains unclear. Is it the learning team's interaction (cooperative learning), the nature of the activity (active learning), or the students' development of mental representations (constructivism) that improves learning outcomes? Or, is it other indirect factors that play an important role, for example, does POGIL encourage students to study more regularly or does it facilitate extracurricular study time with peers (14)? The difficulty of separating out which factors contribute to the long-term retention of information is common in educational studies that adopt a cross-sectional design from necessity (2), making direct comparisons within a single cohort less common.

The Present Study

Brown (2) has identified that learning anatomy and physiology is well suited to the POGIL methodology, but comparisons in his study were made across different year cohorts, and the class sizes were small (with the largest at $n = 31$ students). Undergraduate physiology is often taught in large classes (e.g., $n = 150$ students at our university) with recitation sessions conducted by graduate students untrained in pedagogy or learning theory. Furthermore, recitation sessions in physiology often include active learning tasks because the subject matter lends itself well to such activities (4, 16). For example, students can learn about differences in vision, hearing, touch, and other senses through activities they perform on themselves and one another. Is POGIL more effective than current active learning tasks? Because moving to POGIL involves considerable change both for faculty members and students, further evidence is needed to conclude this methodology produces better long-term retention of information than the current active learning tasks.

We compared performance in an active learning recitation session on physiology of the brain with three variations of POGIL in a quasiexperimental design. In *week 5* of a first-year, first-semester psychology course, recitation sessions were randomly allocated as follows: traditional teaching with active learning tasks (control group); POGIL sessions with an experienced facilitator using the report out component (POGIL group); POGIL sessions without the report out component, using the same facilitator [no report out (NRO) group]; or POGIL sessions run by a graduate student given a brief explanation of POGIL (20 min) but with no experience in POGIL either as a student or as a facilitator [new facilitator (NF) group]. Session allocation to condition was done to ensure each condition had a recitation session run in the morning, afternoon, and evening. The remaining sessions were allocated to conditions to ensure the number of participants in each condition was as equal as possible.

We hypothesised that the POGIL group would show better long-term retention of the material than the control group due to deeper learning and student construction of their own knowledge. As previously stated, the literature suggests that the report out component of POGIL is important to the learning outcomes. We would expect the POGIL group to outperform the NRO group due to the lack of cross-group discussion and justification of their new knowledge. However, this suggestion has never been tested, and it is possible that other intrinsic

components of POGIL provide the learning benefit, in which case we would expect little difference between the POGIL and NRO groups. As with any teaching method, practice improves delivery and performance. T. Vanags has some experience facilitating POGIL activities in her own classes, has reviewed the POGIL literature, and has visited and watched POGIL classes at Seattle University, Franklin and Marshall College, Moravian College, and Virginia Commonwealth University. If experience with POGIL is an important factor in the learning outcomes of students, we expect that participants in the POGIL group would outperform those in the NF group due to the inexperience of graduate student facilitators. If POGIL relies to the greatest extent on the efforts of students (rather than the role of the facilitator), we would expect the NF group to perform on par with the POGIL group.

Finally, evaluations of student learning should include a self-assessment component as this allows students to reflect on their learning (10). Reflecting on the answers they have given and whether they are correct allows students to engage in metacognition, which contributes to recognizing when they understand a concept and when they do not (9). POGIL encourages this self-reflection and assessment, but it is also useful to obtain individual confidence ratings from students in their knowledge both before and after POGIL activities. POGIL facilitators have reported that students feel uncomfortable with the methodology early on (V. Minderhout, personal communication), and it would be helpful to know if their judgements of learning match their learning performance. We hypothesized that participants in the POGIL, NRO, and NF groups would show lower confidence levels in their knowledge after the session activities due to their unfamiliarity with this teaching method.

METHODS

Participants

Participants were 354 undergraduate students enrolled in a first-year introductory psychology unit at an Australian University. Although sexes and ages for the different conditions were not collected, the demographics of the psychology unit overall were 68% women and 32% men with ages ranging from 17 to 54 yr (mean: 20.2 yr, SD: 3.47 yr). Participants did not receive any incentive for participating in this study. The study was run in *week 5* of the semester across the 22 scheduled recitation sessions for that week. Of the original 354 participants, 348 participants completed a posttest quiz at the end of the recitation session and 316 participants completed the (surprise) followup test 2 wk later (see Table 1). This study was approved by the university's Human Ethics Committee.

Table 1. *Numbers of students in each condition at pretest, posttest, and followup*

Time	Control Group	POGIL Group	NRO Group	NF Group	Total
Pretest	94	83	87	90	354
Posttest	93	81	85	89	348
Followup	85	75	79	77	316

Students were assigned to one of the following groups: traditional teaching with active learning tasks (control group); process-oriented guided-inquiry learning (POGIL) sessions with an experienced facilitator using the report out component (POGIL group); POGIL sessions without the report out component, using the same facilitator [no report out (NRO) group]; or POGIL sessions run by a graduate student given a brief explanation of POGIL (20 min) but with no experience in POGIL either as a student or as a facilitator [new facilitator (NF) group].

MATERIALS

Pretest, posttest, and followup quizzes. Quizzes consisted of five free recall questions, such as "Which lobe is responsible for vision?" (answer: the occipital lobe) and "After a brain injury a person is unable to plan. Which lobe is likely to be damaged?" (answer: the frontal lobe). There was a pool of 15 questions for the quizzes with 3 variations of each question. The three variations were constructed to have similar difficulty. For example, *question 1* was one of the following: "Which lobe is responsible for vision?" (answer: the occipital lobe), "Which lobe is responsible for hearing?" (answer: the temporal lobe), and "Which lobe is responsible for memory?" (answer: the frontal lobe). The final part of the pretest quiz was a scale for participants to rate how confident they were that they had answered the questions correctly. The response scale was as follows: *1* = not at all confident, *2* = slightly confident, *3* = neutral, *4* = moderately confident, and *5* = very confident. The pretest, posttest, and followup quizzes were the same, but there were five versions of these quizzes. All five versions were used in all sessions, and the version that a participant received was random.

POGIL/NRO/NF activity sheets. The activity sheets for the POGIL, NRO, and NF conditions consisted of diagrams and questions in the form of four activities on the occipital lobes, parietal lobes, temporal lobes, and frontal lobes. The diagrams were images adapted from pages 95, 97, and 113 of Pinel and Edwards (19). Each diagram was labeled with a moderate amount of information on the function of each brain structure (see APPENDIX A for the frontal lobe activity). There were between three and seven questions associated with each activity. Questions were constructed so that initial questions were directed, and later questions encouraged a deeper understanding of the material (20). Questions could be answered from the diagrams and accompanying information. In keeping with the learning cycle adopted in POGIL activities (exploration, concept formation/term introduction, and application), the name of the brain structure being explored was not given at the top of the page. Participants identified the brain structure in one of the activity questions.

Control activity sheets and resources. The activity sheets for the control condition were similar to the POGIL activity sheets but did not include diagrams. In place of the diagrams and associated information, a variety of textbooks, reference books, and three-dimensional brain models were provided. Finally, unlike the POGIL activities, each control activity was labeled with the name of the brain structure, for example, *Activity 4: Frontal Lobe* (see APPENDIX B for the control group frontal lobe activity).

Manager instructions and team role descriptions. Participants in all conditions except the control condition were given a one-page handout with a description of the team roles of manager, presenter, recorder, and reflector/strategist based on the descriptions in Hale and Mullen (8). The manager was assigned by the facilitator, and the manager allocated the role of recorder to the person on their left, the role of presenter to the person on their right, and the role of reflector/strategist to the remaining person (for those in groups of 4). The manager's instruction sheet was based on the materials provided in a POGIL workshop run by Prof. Minderhout ("I Already Teach Well, Why Should I Change? Active Learning in Large Lectures" Seminar at Australian National University on July 12, 2010). There were two rules added to this information. The first rule stated that if group members had questions, they should discuss the questions as a group first, and if they could not resolve the questions, the manager should put the questions to the facilitator. The second rule was that as a group they should come up with one set of answers for the activity (rather than individual answers being recorded on individual sheets).

Experimental Design

The experiment was a four (control, POGIL, NRO, NF) by three (pretest, posttest, followup) mixed factorial design with one between-

subjects factor (teaching method) and one within-subjects factor (time of recall).

Procedures

Facilitators. The POGIL facilitator for this experiment was T. Vanags. As T. Vanags works at a different university, she was unknown to the students. She was introduced as a researcher in the area of teaching and learning and had attended the first lecture of the semester advising students that some evaluations of teaching methodologies would take place during the semester. T. Vanags was experienced with POGIL, having used POGIL learning teams and activities in her own classes and having reviewed a proportion of the literature before undertaking a study tour of several universities using POGIL in the United States.

T. Vanags ran the POGIL and NRO sessions with the regular tutor in attendance but uninvolved. Control sessions were run by the regular tutors for those sessions. These were all graduate students enrolled in either an Honours or a PhD program. NF sessions were run by the sessions' regular tutors. These new facilitators had no experience of POGIL either as a student or as a facilitator but had been given a 20-min briefing and given the Farrell et al. (6) article to read before taking the session (see Table 2 for a breakdown of conditions and facilitators). The study was run in 22 recitation sessions across a single week, and T. Vanags was present in all sessions to minimize differences across groups that may be attributed to the presence of an unknown authority figure.

At the start of the study, participants completed the pretest quiz. Participants expected this quiz as it was an assessable component (5%) of their course. It was consistent in terms of content and layout to five other recitation assessable quizzes during the semester. The reading for this quiz was pages 99–113 of the textbook *Psychology* (3). Participants had 10 min to answer the questions on the quiz sheet before handing it in.

Control groups. The tutor assigned participants to teams of four where possible, with one or two groups of three participants in some sessions. No roles were given to group members and no rules were set for groups. Participants were then given the handouts and asked to use the various reference books, textbooks, and three-dimensional brain models to work as a group in answering the questions. When the tutor saw that participants had answered all the questions, s/he asked each group to verbally provide the answer to one or more questions for the whole class and provided corrections where appropriate. Participants worked on *activities 1* and *2* together and then on *activities 3* and *4*. In all conditions, the two sets of activities took ~30 min each. Once all four activities were complete (~30 min before the session ended), participants were asked to put away their activity sheets and were given the posttest quiz. Participants had 5 min to complete the posttest quiz. Quizzes were collected, and this marked the end of the experiment.

Experimental groups (POGIL/NRO/NF). As with the control groups, the facilitators assigned participants to teams of four where possible and teams of three otherwise. The facilitator then gave students a brief overview of the POGIL methodology (based on the team role description sheet). Once groups were formed and participants had read the team role descriptions and manager's instructions, participants were given the POGIL activity handout. Participants were asked to work as a team and adopt their roles within the team as they completed the activities.

POGIL AND NF CONDITIONS. As participants completed the questions in the first two activities, the facilitator asked the groups' presenters to come to the front of the classroom and write their answers to one question on the whiteboard. When all questions had been answered on the whiteboard, the facilitator reviewed the answers and asked everyone for comments or amendments. As with the control condition, once all four activities were completed, participants were asked to put their activity sheets away and they had 5 min to complete the posttest quiz.

NRO CONDITION. This condition was identical to the POGIL and NF conditions described above; however, students were not asked to write their answers on the whiteboard. Instead, the facilitator asked each question aloud and waited for teams to volunteer the answer. If teams answered and the answer was incorrect or incomplete, the facilitator provided the correction and/or additional information. If the teams did not answer, the facilitator provided the answer. As with the other conditions, after completing all four activities, participants put away their activity sheets away and completed the posttest quiz.

Followup. In the recitation sessions 2 wk after the experiment, all participants were given 5 min to complete the followup quiz. Participants did not know that they would be given this quiz and their results would not be included in their final grade.

RESULTS

Pretest Knowledge

To ensure the knowledge of the participants in the four conditions was comparable before participation in the study, one-way ANOVA was done on students' pretest knowledge. The pretest quiz was an assessable quiz participants were expecting as part of the course. Unexpectedly, control group participants had higher pretest scores than participants in the other three groups [$F_{(3,350)} = 5.956$, $P = 0.001$; see Fig. 1]. Post hoc tests showed that the control group outperformed the POGIL group [$t_{(175)} = 2.805$, $P = 0.006$], the NRO group [$t_{(179)} = 3.092$, $P = 0.002$], and the NF group [$t_{(182)} = 4.264$, $P < 0.0005$]. There were no significant differences between the POGIL, NRO, or NF groups.

Examination of Differences in Pretest Knowledge

As we did not expect a difference in pretest knowledge across groups, we carried out further analyses in an attempt to identify the factor(s) involved. Claims are often made that the time of day affects student engagement and performance, but one-way ANOVA revealed no differences in performance across morning, afternoon, or evening recitation sessions ($P = 0.14$).

We then conducted one-way ANOVA with recitation session number as the independent variable (students choose a recitation session before the semester starts and remain in that session for the semester). There was a strong effect for recitation session number [$F_{(21,332)} = 2.600$, $P < 0.0005$; see Fig. 2].

Table 2. *Facilitators and those present in the recitation sessions for each condition*

	Control Group	POGIL Group	NRO Group	NF Group
Facilitated by	Regular tutor	T. Vanags	T. Vanags	Regular tutor
POGIL expertise	None	Moderate	Moderate	20-min briefing
Present in class	Regular tutor and T. Vanags	Regular tutor and T. Vanags	Regular tutor and T. Vanags	Regular tutor and T. Vanags

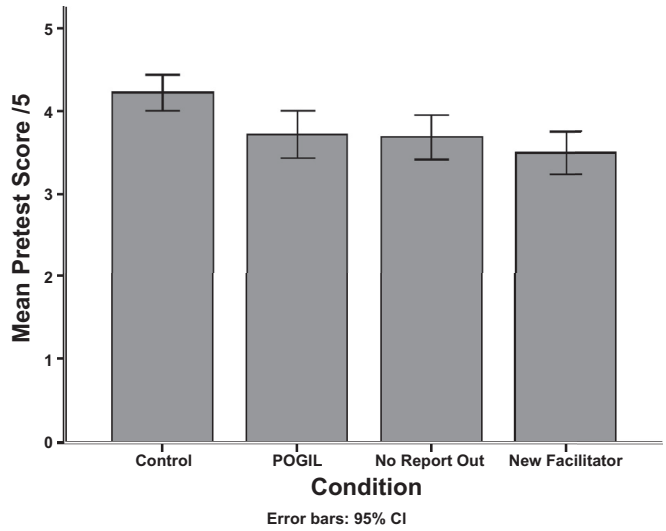

Fig. 1. Mean scores for the following conditions: traditional teaching with active learning tasks (control group; $n = 94$); process-oriented guided-inquiry learning (POGIL) sessions with an experienced facilitator using the report out component (POGIL group; $n = 83$); POGIL sessions without the report out component, using the same facilitator [no report out (NRO) group; $n = 87$]; and POGIL sessions run by a graduate student given a brief explanation of POGIL (20 min) but with no experience in POGIL either as a student or as a facilitator [new facilitator (NF) group; $n = 90$]. Control group participants had higher pretest scores than participants in the other three groups, and there were no significant differences across the POGIL, NRO, or NF groups. Error bars = 95% confidence intervals.

The large number of sessions ($n = 22$) ruled out the possibility of conducting post hoc tests due to the chance of a type I error. We suggest that the pretest differences across recitation sessions are the result of differing abilities and motivations of students within different sessions. For example, *session 1* is the first class of the week at 9 AM on Monday mornings. It may be that students selecting this class are high achieving students who are keen to learn. The pretest results allowed us to use recitation number as a covariate in the remaining analyses, thereby accounting for preexisting differences across class groups.

Retention of Knowledge Over Time for Different Teaching Methods

A four (control, POGIL, NRO, NF) by three (pretest, posttest, followup) split plot analysis of covariance (ANCOVA; covariate: recitation number) showed a main effect for time [Wilks' $\Lambda = 0.980$, $F_{(2,308)} = 3.080$, $P = 0.05$, multivariate partial $\eta^2 = 0.020$] and a main effect for teaching method [$F_{(3,309)} = 3.478$, $P = 0.02$, multivariate partial $\eta^2 = 0.033$], but no time by teaching method or time by recitation number interactions ($P = 0.07$ and $P = 0.99$, respectively; see Fig. 3).

Effect of teaching method (between subjects). Immediately after the teaching activities (posttest), the POGIL group performed more poorly than the control group [$t_{(172)} = 2.237$, $P = 0.03$]; however, there were no differences in performance between the POGIL and NRO groups ($P = 0.22$) or between the POGIL and NF groups ($P = 0.09$) at this time. There were no differences in performance between the control and NRO ($P = 0.31$) or NF ($P = 0.53$) groups at posttest. At followup 2 wk later, the POGIL group performed on par with the control group ($P = 0.35$) and the NRO ($P = 0.87$) and NF ($P = 0.61$) groups.

Effect of time (within subjects). CONTROL GROUP. The control group did not improve as a result of the active learning (non-POGIL) activities in the recitation session ($P = 0.84$). Furthermore, they failed to retain the knowledge they brought to this session over the next 2 wk, performing more poorly at followup than they did on the posttest [$t_{(84)} = 4.278$, $P < 0.0005$]. Importantly, their knowledge at followup was significantly less than their pretest knowledge when they had studied for the assessable quiz [$t_{(84)} = 3.830$, $P < 0.0005$; see Table 3].

POGIL GROUP. Participants in the POGIL group did not show any improvement immediately after the POGIL activities ($P =$

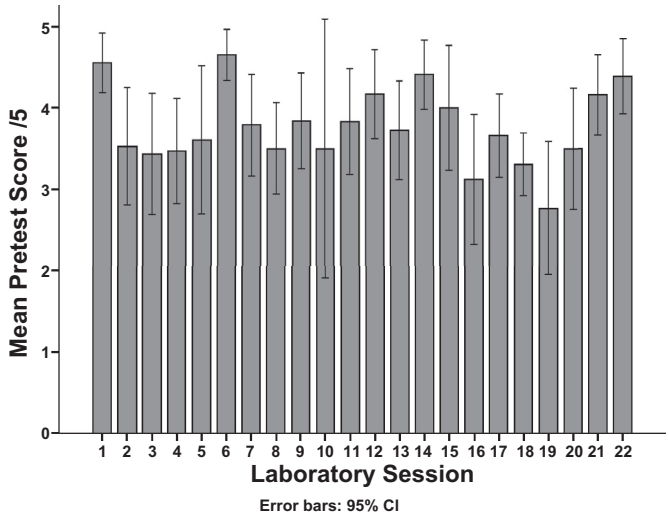

Fig. 2. Mean scores on the pretest quiz for all 22 recitation sessions. Participant numbers in each session were as follows: $n = 20$ (*sessions 6*, *13*, and *16*), $n = 19$ (*sessions 1*, *7–9*, *17*, and *21*), $n = 18$ (*sessions 11* and *12*), $n = 17$ (*session 4*), $n = 16$ (*sessions 2*, *3*, and *14*), $n = 15$ (*session 19*), $n = 14$ (*session 5*), $n = 13$ (*session 18*), $n = 12$ (*sessions 15* and *20*), $n = 9$ (*session 22*), and $n = 4$ (*session 10*). Error bars = 95% confidence intervals.

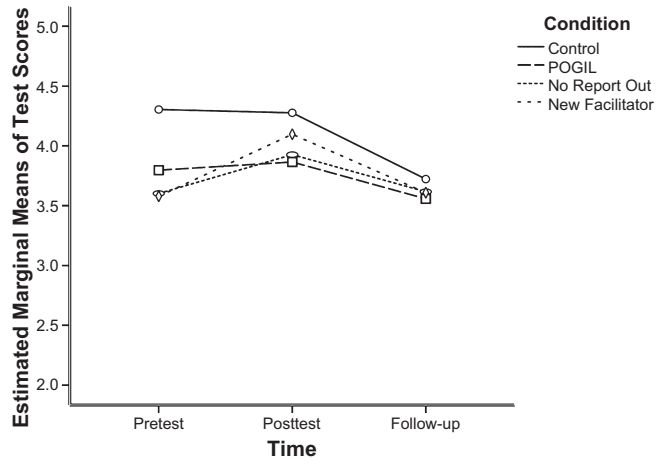

Fig. 3. Mean scores for the control, POGIL, NRO, and NF groups on the pretest, posttest, and followup quizzes. Participants in the control group had higher pretest scores than participants in the other three groups, and the POGIL group performed more poorly than the control group at posttest. At followup, there were no differences across groups; however, the control group knew significantly less than at pretest. The POGIL, NRO, and NF groups performed on par with their pretest scores.

Table 3. *Mean number of questions correct (out of 5 questions total) for each condition at pretest, posttest, and followup*

Time	Control Group	POGIL Group	NRO Group	NF Group
Pretest	4.3 (1.1)	3.8 (1.3)	3.7 (1.3)	3.6 (1.2)
Posttest	4.2 (0.9)	3.8 (1.1)	4.0 (1.0)	4.1 (0.9)
Followup	3.7 (1.1)	3.5 (1.2)	3.7 (1.2)	3.6 (1.2)

Values are means (SD).

0.54), and the drop in their knowledge over the next 2 wk almost reached significance ($P = 0.05$). However, a comparison of their pretest and followup results showed that, unlike the control group, they retained the information they had learned in preparation for the assessable quiz. There were no differences in their pretest and followup performance ($P = 0.16$).

NRO GROUP. Participants in the NRO group improved as a result of the POGIL activities [$t_{(84)} = 3.150$, $P = 0.002$], but these participants forgot some of that material over the next 2 wk [$t_{(77)} = 2.162$, $P = 0.03$]. A comparison of their pretest and followup performance showed that the NRO group retained the knowledge they brought to the assessable quiz ($P = 0.94$) but not the knowledge they had gained in the POGIL session.

NF GROUP. Similarly to the NRO group, participants in the NF group showed an improvement immediately after the POGIL activities [$t_{(88)} = 5.129$, $P < 0.0005$] but forgot some of that material between posttest and followup [$t_{(75)} = 3.675$, $P < 0.0005$]. As with the other POGIL groups, their performance at followup was no different than their performance at pretest ($P = 0.94$), indicating that pretest levels were retained over a 2-wk period.

Confidence

Using one-way ANOVA, we evaluated the confidence of participants in their pretest, posttest, and followup knowledge. There were no differences in confidence levels across groups at pretest ($P = 0.15$), but the differences at posttest [$F_{(3,335)} = 4.246$, $P = 0.006$] and followup [$F_{(3,385)} = 5.535$, $P = 0.001$] were significant. Immediately after the class activities, POGIL participants had significantly lower levels of confidence in their knowledge than those in the control ($P = 0.04$) and NRO ($P = 0.006$) conditions. Although the POGIL group had lower levels of confidence than the NF group, this result only approached significance ($P = 0.06$). Two weeks later, the POGIL group remained less confident than the control ($P = 0.03$), NRO ($P = 0.002$), and NF ($P = 0.004$) groups.

DISCUSSION

In this study, we compared the learning outcomes from POGIL sessions with current active learning methods (taking apart plastic models of the brain). We investigated whether POGIL leads to better long-term retention of information than traditional active learning tasks. The first hypothesis that participants in POGIL sessions would show better long-term retention of material than those in control groups was supported as none of the POGIL groups showed the same memory decay of knowledge as the control group over the subsequent 2 wk. At followup, the control group performed more poorly than they did on the pretest quiz. The pretest quiz was an assessable quiz they expected and for which they had studied [the material covered in the reading assignment (for the pretest quiz) was the textbook content on the topic covered in the previous lecture and in the recitation sessions of this research]. In contrast, the POGIL groups did not show any decrement in performance between pretest and followup, indicating that they had retained the material they had studied.

We also investigated whether the report out component of POGIL was critical to student learning. The second hypothesis that the POGIL group would outperform the NRO group due to a lack of cross-group discussion in the NRO condition was not supported. In fact, although there were no differences in POGIL and NRO group performances at posttest, the NRO group did show an improvement in knowledge immediately after the POGIL activities, which the POGIL group did not. Feedback from a small number of participants in the POGIL group indicated that they began to question their existing knowledge after the POGIL activities. This may have been due to students rote learning the information for the pretest quiz [many quizzes had acronyms on them, such as FPOT (frontal, parietal, occipital, temporal) for the 4 lobes]. As the POGIL activities were only 1 h in total, it is possible that students had time to begin developing a deeper understanding of the brain anatomy but insufficient time to master that information, leading to a loss of the surface learning in which they had previously engaged. The additional knowledge gained by the NRO group was not retained over the next 2 wk; however, as with the POGIL group, the long-term retention of material was retained over the next 2 wk.

Another component investigated was the importance of teaching assistant training when implementing POGIL. We hypothesized that if facilitator experience with POGIL is an important factor in learning outcomes for students, students in the NF group would perform more poorly than those in the POGIL group. This hypothesis was not supported. As with the NRO group, NF participants actually gained additional knowledge as a result of the POGIL activities but similarly failed to maintain that additional knowledge over the next 2 wk. Despite this, as with the POGIL and NRO groups, participants in the NF condition retained their pretest knowledge over the 2-wk period between the posttest and followup.

The findings that all POGIL groups retained the material over a 2-wk period is particularly impressive for two additional reasons: *1*) experienced POGIL users report that many POGIL implementers find student performance can be adversely at the introduction of POGIL because the methodology is new to students and they must come to terms with learning in a "new" manner and the expectations associated with this new style of interaction (R. S. Moog, personal communication); and *2*) the students in the POGIL and NRO groups had a stranger as a facilitator, and all students had the "stranger" in their recitation session, which may have raised students' levels of anxiety or affected their level of engagement. Indeed, the weaker posttest results of the POGIL group may have been due to the fact that this group had a stranger as a facilitator and they were accountable for generating their answers due to the report out component. The final hypothesis that students in the POGIL groups would show lower levels of confidence in their knowledge than those in the control group was partially supported. The POGIL group had lower levels of confidence than the control group, but the NRO and NF groups did not. There were no differences

in the confidence levels of the control, NRO, and NF groups at posttest or followup. The POGIL group showed lower levels of confidence than the other groups at both time points.

The findings from this study are interesting for several reasons. First, although the quasiexperimental design of allocating different experimental interventions to different recitation groups is common in educational research in psychology, it is less common to evaluate pretest knowledge differences across these groups. Our results indicate these pretest measures are vital as homogeneity of recitation groups cannot be assumed. As teachers, we have all experienced variation in class performance across a cohort, and failing to obtain pretest evaluations showing such variation can lead to erroneous conclusions about the success or failure of a teaching trial (15). It is possible, as occurred in this study, that the majority of the higher (or lower) performing participants will be clustered within one experimental condition.

Perhaps the most interesting aspect of this study is that despite the control group showing the best performance at pretest, this group failed to retain that knowledge long term. Although this group had clearly studied hard and effectively for their assessable quiz, that knowledge was subject to normal memory decay with time. In contrast, all POGIL groups (POGIL, NRO, and NF) retained the knowledge they had acquired for the quiz over the 2-wk period. This suggests that POGIL activities, regardless of whether they are run by trained or untrained facilitators and regardless of the report out component, produce the deeper learning that results from students constructing (or reconstructing) knowledge themselves. This is a different type of learning to the rote learning most students engage in when preparing for a quiz or exam when they are learning to pass the exam rather than to understand the material (1, 21).

With regard to performance immediately after the activities (posttest), we were surprised that the students in the NRO and NF groups performed better than the POGIL group. First, the POGIL and NRO groups were both facilitated by T. Vanags, suggesting the recitation sessions should have produced similar learning outcomes except that the NRO group lacked the opportunity for cross-group discussion and the associated deeper learning. Second, the NF groups were run by the teaching assistants inexperienced with POGIL, and we expected experience with a teaching methodology to improve learning outcomes. However, both assumptions appear incorrect. It might be suggested that omitting the report out function improves immediate performance, but the lack of difference between the NRO and NF groups immediately after the activity does not support this contention (the NF group did report out). Another assumption could be that the teaching assistants provided better instruction to students than T. Vanags; however, T. Vanags' student satisfaction ratings over her career and her multiple teaching awards would suggest this is unlikely. It is possible that as the NF groups' facilitators were their regular tutors they may have been more comfortable with those individuals than with T. Vanags. The regular tutors may have been considered less of an authority figure. However, there were no differences in posttest performance of the NRO (T. Vanags) and NF (regular tutor) groups, suggesting that instruction ability and familiarity of the facilitators were not factors. Another consideration could be that the new facilitators were closer in age and recent undergraduate experience to their participants and may have better understood which concepts were difficult to master.

Considering the results of the POGIL, NRO, and NF groups in conjunction with confidence ratings and our reflections on the study, it seems to the authors that the role of affirming knowledge may be important in POGIL implementation. T. Vanags adhered closely to the recommendations in the POGIL literature avoiding confirming or disconfirming answers (21). However, she only adopted this approach in the POGIL group. In the NRO condition, when she asked presenters to volunteer answers (as there was no report out), she felt it necessary to provide answers and explanations when no responses were forthcoming. In addition, because there was no information written on the board for students to refer to during discussions, she gave more direct responses to the POGIL questions from the worksheets. As she was present in all recitation sessions, T. Vanags was able to compare this with how other sessions were run. The observations of T. Vanags are that the new facilitators (as well as the control group facilitators) were very direct in their responses to the presenters' answers. They would confirm or disconfirm answers given by students, most likely because this is the way teaching normally occurs in classrooms. In other words, the POGIL group were the only group left to construct the correct answers by intra- and cross-group discussion and argument, and this group did not receive definitive statements about the accuracy of their answers. While T. Vanags ensured the information agreed upon by the groups was correct and complete, but she did not vocalize this to the participants. Participants may have felt unsure about their knowledge after discussions and disagreements without a summation by an authority figure that "this information represents the correct answer." This is also likely to have affected the participants' confidence in their knowledge, resulting in the lower confidence scores we predicted POGIL participants would show.

The attempt in this study to breakdown POGIL and identify crucial elements is not without limitations. The results suggest that POGIL and variations of POGIL teaching produce better long-term retention of information than the existing active learning tasks; however, the evaluation of POGIL was based on only one recitation session. POGIL requires participants to take on specific roles within teams and to take on different roles at different times. The literature suggests this is an important facet of POGIL, and it requires POGIL implementation over a longer timeframe. In this study, there was no rotation of team member roles, and students are unlikely to have fully embraced their roles because of the limited time to become familiar with the roles and methodology. Future research could evaluate participant performance on a weekly basis. This would help to determine whether rotating the roles and becoming familiar and confident with POGIL affects learning outcomes.

Another limitation of this study is that the materials used for the POGIL activity did not extend students a great deal in their understanding of the material (see APPENDIX A). These were the first POGIL activities developed by these authors, and they are likely to be unsophisticated. In addition, the actual POGIL implementation in this study is not necessarily the way others would implement POGIL. Indeed, the value of POGIL is that it can be implemented in many ways, and conclusions drawn in this study should be tested by other POGIL proponents using their own approaches to see if they can be replicated. The study does provide confirmation of anecdotal evidence that POGIL initially

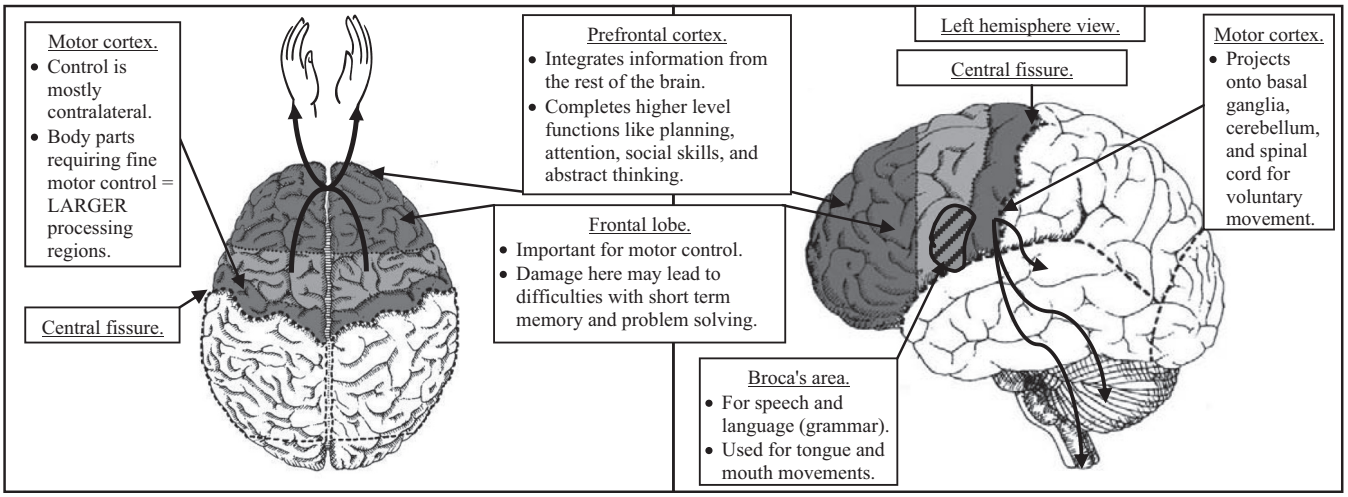

Fig. 4. Diagram adapted from Ref. 19 (p. 113).

decreases students' confidence levels (V. Minderhout, personal communication), although it must be remembered that in the POGIL condition, students did not experience the "closure" that would normally be used in POGIL and reflects best practice in helping students link classroom activities and learning outcomes. Even so, it would also be valuable to collect individual confidence ratings at different time points to see whether participant confidence increases with exposure to POGIL.

This study demonstrates that POGIL approaches to teaching can produce improved long-term learning outcomes for students even when the teaching is done by less experienced teaching assistants and when the report out component is omitted, an important factor for large lecture classes in which reporting out may be logistically impractical. The results also suggest that to become confident in their knowledge, students must not only construct their knowledge through grappling with the material but must also receive clear indications that their constructed knowledge is valid. While student-cenetred learning is key to engaging students in their studies and allowing them to take responsibility for their learning, constructing their own knowledge and assurance from facilitators that these mental representations are accurate are fundamental to long-term recall and confidence in that knowledge.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: T.V., K.P., and J.B. conception and design of research; T.V. performed experiments; T.V. analyzed data; T.V., K.P., and J.B. interpreted results of experiments; T.V. prepared figures; T.V. drafted manuscript; T.V., K.P., and J.B. edited and revised manuscript; T.V., K.P., and J.B. approved final version of manuscript.

APPENDIX A: POGIL/NRO/NF ACTIVITY (FRONTAL LOBE)

Activity 4

Examine the diagram of the brain (shown in Fig. 4).

1. Mark the front and back of the brain and the occipital and parietal lobes on the diagram. What structure separates the shaded area from the parietal lobe?

2. We saw that the parietal lobe contains the somatosensory cortex, which processes sensory information from the body. What is the structure in the darker shading just in front of the parietal lobe and what is its function?

3. What do we call the area at the base of this darkly shaded area that is responsible for speech production? Why do you think this is key to speech *production* rather than speech comprehension (hint: think about the word *production*–how do we make sounds)?

4. What do we call the darkly shaded part at the front the brain? What are some of the functions of this area?

APPENDIX B: CONTROL ACTIVITY (FRONTAL LOBE)

Activity 4: Frontal Lobe

1. What structure separates the frontal lobe from the parietal lobe?

2. We saw that the parietal lobe contains the somatosensory cortex, which processes sensory information from the body. What is the similar structure in the frontal lobe and what does it do?

3. In the frontal lobe, there is a section called Broca's area. What is the role of this structure? What would happen if this structure was damaged?

REFERENCES

1. **Biggs J, Tang C.** *Teaching for Quality Learning at University*. New York: Open Univ. Press McGraw Hill Education, 2009.
2. **Brown PJ.** Process-oriented guided-inquiry learning in an introductory anatomy and physiology course with a diverse student population. *Adv Physiol Educ* 34: 150–155, 2010.
3. **Burton L, Westen D, Kowlaski R.** *Psychology: Australian and New Zealand Edition*. Queensland: Wiley and Sons, 2008.
4. **Carvalho H, West CA.** Voluntary participation in an active learning exercise leads to a better understanding of physiology. *Adv Physiol Educ* 35: 53–58, 2011.
5. **Eberlein T, Kampmeier J, Minderhout V, Moog RS, Platt T, Varma-Nelson P, White HB.** Pedagogies of engagement in science. *Biochem Mol Biol Educ* 36: 262–273, 2008.
6. **Farrell JJ, Moog RS, Spencer JN.** A guided inquiry general chemistry course. *J Chem Educ* 76: 570–574, 1999.
7. **Gillespie RJ, Spencer JN, Moog RS.** Demystifying introductory chemistry. *J Chem Educ* 7: 617–622, 1996.
8. **Hale D, Mullen LG.** Designing process-oriented guided-inquiry activities: a new innovation for marketing classes. *Market Educ Rev* 19: 73–80, 2009.
9. **Hanson DM.** *Instructor's Guide to Process-Oriented Guided-Inquiry Learning*. New York: Pacific Crest, 2006.

10. **Hanson DM, Wolfskill T.** Process workshops–a new model for instruction. *J Chem Educ* 77: 120–130, 2000.
11. **Hardy I, Smith E.** Contesting tertiary teaching qualifications: an Australian perspective. *Teach Higher Educ* 11: 337–350, 2007.
12. **Hinde RJ, Kovac J.** Student active learning methods in physical chemistry. *J Chem Educ* 78: 93–99, 2001.
13. **Landis CR, Peace GE, Jr, Scharberg MA, Branz S, Spencer JN, Ricci RW, Zumdahl SA, Shaw D.** The new traditions consortium: shifting from a faculty-centred paradigm to a student-centred paradigm. *J Chem Educ* 75: 741–744, 1998.
14. **Lewis SE, Lewis JE.** Departing from lectures: an evaluation of a peer-led guided inquiry alternative. *Chem Educ Res* 82: 135–139, 2005.
15. **Lewis SE, Lewis JE.** Seeking effectiveness and equity in a large college chemistry course: an HLM investigation of peer-led guided inquiry. *J Res Sci Teach* 45: 794–811, 2008.
16. **Michael J.** Where's the evidence that active learning works? *Adv Physiol Educ* 30: 159–167, 2006.
17. **Moog RS, Creegan FJ, Hanson DM, Spencer JN, Straumanis AR.** Process-oriented guided inquiry learning: POGIL and the POGIL project. *Metropol Univ J* 17: 41–51, 2006.
18. **Moog RS, Spencer JN.** (editors). *Process-Oriented Guided Inquiry Learning*. Washington, DC: Oxford Univ. Press, 2008.
19. **Pinel JP, Edwards M.** *A Colorful Introduction to the Anatomy of the Human Brain: a Brain and Psychology Coloring Book*. Boston, MA: Pearson Education, 2008.
20. www.pogil.org. *Elements of a Typical POGIL Classroom Activity* (online). http://www.pogil.org/uploads/media_items/rb9-jun-10-elements-of-pogilactivity.original.pdf [10 July 2013].
21. **Spencer JN.** New directions in teaching chemistry: a philosophical and pedagogical basis. *J Chem Educ* 76: 566–569, 1999.