Weidner, J., & Ranney, M. (2000). Knowledge retention following problem-solving versus information-gathering. In R. Robson (Ed.) *Proceedings of the 2000 International Symposium on Mathematics/Science Education and Technology* (pp. 404-409), Charlottesville, VA: AACE.

Knowledge Retention Following Problem-Solving Versus Information Gathering

Jeanne Weidner and Michael Ranney
Graduate School of Education
University of California, Berkeley
United States
jweidner@socrates.berkeley.edu and ranney@soe.berkeley.edu

Abstract: This study investigated the retention of concepts and knowledge organizations six months after an initial phase during which subject pairs used computer technology to support two divergent instructional goals: (a) the solving of a clinical problem versus (b) gathering factual information to answer direct questions. After the intervention, the information gathering activity yielded significantly higher performance on the outcome measures (e.g., gain scores, post-tests and PFNET correlations) compared to the problem solving activity. However, this advantage disappeared upon delayed testing six months later, as the information gathering context yielded significant declines on all measures, while there were no such declines regarding the problem solving context. In addition, heterogenous academic pairs and homogenous gender pairs exhibited superior performance on initial testing, a finding that persisted to some degree upon delayed testing.

Introduction

Many of the cognitive science applications for instructional practice have emphasized ways to make classroom science learning a more active process. One salient approach has been problem based learning (PBL), which generally begins with a problem, and requires students to acquire concepts and facts that will ultimately assist in solving it. Two advantages credited to this method have been (a) a greater retention of knowledge and (b) an increased ability to apply it (Eisenstaedt, Barry, & Glanz, 1990; Norman & Schmidt, 1992; Albanese & Mitchell, 1993). Support for these claims may lie in current cognitive theory which would seem to suggest that more durable cognitive structures are developed during PBL instruction. While rather non-contextualized learning has a greater tendency to develop inert knowledge that is stored in memory without any indication of how it will be used, problem-based learning develops knowledge that must be used to solve a specific problem and is thought to be incorporated into cognitive structures that may be well-modeled as production systems (Anderson, 1987). By applying knowledge in an active manner, the learner is thought to develop production systems in which the knowledge becomes encoded into more robust and longer-lasting frameworks. For example, it is possible that problem solving affects the way that knowledge is organized and accessed, giving rise to cognitive structures that account for some of the benefits often attributed to PBL. If problem solving results in different knowledge organizations than those emerging from other activities (such as responding to questions or gathering information), then these differences should be able to be captured and compared by measures such as semantic networks.

We chose Pathfinder Networks (PFNETs; Schvaneveldt, 1990) as one of our outcome measures because they have been shown to discriminate among the knowledge organizations of subjects who use information for divergent purposes. Each network representation (PFNET) yields a two-dimensional "concept map" of a subject's knowledge organization, based on his/her subjective rating of the relatedness of pairs of terms (as in Figure 1). For example, Durso, Rea, and Dayton (1994) used Pathfinder Associative Networks to measure the knowledge organizations of subjects solving an insight problem, relative to those who were given the information as a story, rather than as a problem to solve. They reported that people who solved an insight problem had a significantly different knowledge organization, as measured by PFNETS, than did those who (a) did not solve it, or (b) were presented with the information in a non-problem format. In this study, we selected pairs of terms related to cranial nerves and utilized Pathfinder Networks to determine whether or not problem-solving yields different knowledge organizations when compared to query-driven information gathering. Other researchers have found that students' performances in a course correspond well with the correlation of their Pathfinder Networks with Pathfinder Networks generated by their instructor (Goldsmith & Johnson, 1990).

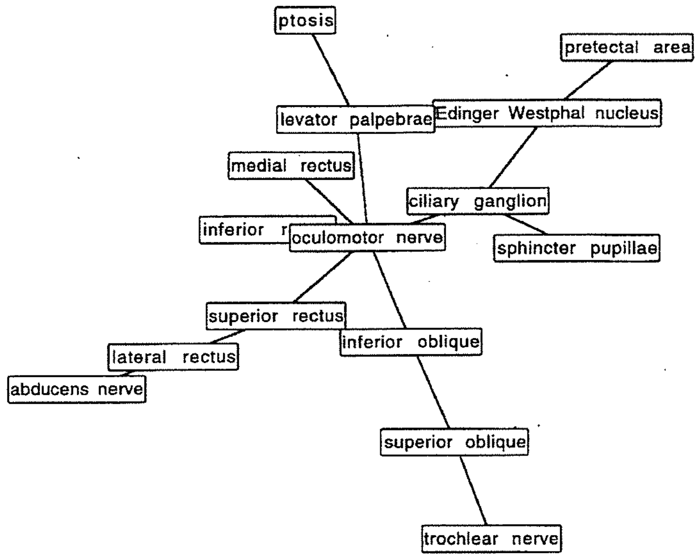

Figure 1: Composite expert PFNET for Exercise 1.

Initial Phase of the Study

As reported previously (Weidner, Ranney & Diamond, 1999), we studied the ways in which pairs of students used the representationally rich multimedia program *BrainStorm: An Interactive Neuroanatomy Atlas* (Coppa & Tancred, 1995) under two contexts – while solving a clinical problem due to cranial nerve dysfunction and while gathering information in order to respond to direct questions about these cranial nerves. For example, the problem solving (PS) context presented the symptoms of a patient suffering from a lesion of the oculomotor nerve, whereas the information gathering (IG) context would ask a question such as "What is the function of the oculomotor nerve?" Each pair of subjects completed two exercises, one in the PS context and one in the IG context. The study was counterbalanced such that the subset of subjects who were assigned a PS context for the first exercise received the IG context for the second, and vice versa. A total of 22 students (13 graduates and 9 undergraduates; 11 males and 11 females) participated in the study.

Assessments completed individually by each subject included a (20- or 23-point) pre-test prior to each exercise, and a 50-point post-test after the exercise (which included the corresponding pre-test items). By comparing the pre-test score to the performance on those corresponding items from the post-test, a *gain score* was calculated for each subject. Subjects' scores on the additional, non-corresponding, items on the post-test (the *supplementary post-test*) were also analyzed as an outcome measure. In addition, the subjects were given a list of 15 terms from each exercise, which was used to generate Pathfinder Networks (PFNETs) after the completion of each exercise. Five experts in the domain also used the same 15 terms/exercise to generate PFNETs in order to determine whether these networks would differentially correlate with subjects' PFNETS as a function of context. A composite of the experts' PFNETs was used for comparison with each subject PFNET yielding a correlation (*PFNET correlation)*.

Findings from the initial phase revealed that, relative to the mean from the problem-solving (PS) context, the information gathering (IG) context yielded marginally higher gain scores (IG = 23.4% vs. PS = 15.5%; t = 1.866,

$p = 0.08$) and its PFNETS were marginally more highly correlated with the experts' ® (IG) = .574 vs. r (PS) = .513; $t = 1.960$, $p = .06$). In addition, since the IG context resulted in higher mean scores for all outcome measures across both exercises, non-parametric tests revealed a significantly higher performance overall for the IG context across these outcome measures (Wilcoxon Signed Ranks Test: $Z = 2.201$, $p < .05$). An additional significant measure of the superior performance under the information gathering context is evidenced by the fact that a greater number of the experts' PFNETs correlated more highly with PFNETs resulting from the IG instructional context, relative to those from the PS context (Wilcoxon Signed Ranks Test: $Z = 2.201$, $p < .05$).

With regard to the effects of pair membership on performance, significant findings from this initial phase showed that undergraduates working in academically heterogenous pairs (i.e., with graduate students), compared to their peers who worked in homogenous pairs (i.e., undergraduates with undergraduates), scored higher on two of the outcome measures (with gain score means of 30% vs. 14%; $F_{(1,17)} = 4.74$, $p < .05$, and PFNET correlations of .610 vs. .371; $F_{(1,17)} = 16.07$, $p < .01$). Graduate students seemed to also benefit from working in a mixed academic pair, as indicated by their marginally higher performance on PFNET correlations (.692 vs. .582; $F_{(1,25)} = 3.54$, $p = .07$). Non-parametric tests comparing means of all measures indicated that both undergraduate and graduate students performed better in mixed academic pairs than their counterparts in homogenous pairs (Wilcoxon Signed Ranks Test: $Z = 2.201$, $p < .05$; cf., White and Frederiksen, 1998, on the mutual benefit of heterogenous ability pairings in younger subjects).

Further results from this initial phase revealed differences based on gender pairings. Non-parametric tests comparing means of all measures indicated that both males and females performed better in same-gender pairs than their counterparts who were in mixed gender pairs (Wilcoxon Signed Ranks Test: $Z = 2.201$, $p < .05$; cf., Underwood, Jindal & Underwood, 1994). Additionally, males who were paired with males rather than females yielded significantly higher gain scores by analysis of variance (26% vs. 15%; $F_{(1, 21)} = 5.01$, $p < .05$).

Delayed Phase of Study

Subjects who completed the initial phase of the study were contacted by email or telephone to participate in a follow-up phase, and were paid for their participation. All 22 subjects who participated in the initial phase were contacted, with only one declining to participate. The delayed phase was designed to compare subjects' performance on the outcome measures (e.g., gain score, supplementary post-test, and PFNETs) six months after the completion of the original exercises. The overall declines between initial and delayed scores for each context, and the difference in absolute decay between contexts were also of interest.

Context	Measure	Initial Post-Test	Delayed Post-Test	Difference	t	p value
Information Gathering	Gain Score (from Pre-test)	0.239	0.157	-0.083	-3.523	0.002
	Supplementary Post-Test Score	0.586	0.525	-0.061	-2.172	0.042
	PFNET correlation	0.583	0.454	-0.130	-3.061	0.002
Problem Solving	Gain Score (from Pre-test)	0.165	0.111	-0.054	-1.565	0.133
	Supplementary Post-Test Score	0.573	0.544	-0.029	-1.385	0.181
	PFNET correlation	0.510	0.485	-0.025	-0.816	0.424

Table 1: Paired t-tests of the differences between delayed and initial outcome measures for both IG and PS contexts.

Findings from this phase showed that, in contrast to the results from the initial post-test, analyses of variance for the delayed post-test revealed no significant or marginal differences whatsoever between the two contexts (i.e., across gain scores, supplementary post-tests, or PFNET correlations). Further, the initially significant non-parametric difference between the two contexts' performance-yields on the various independent measures was absent during the delayed testing. As expected, both contexts yielded an overall decline for all measures from initial post-test scores to delayed post-test scores. However, paired t-tests indicated that while these declines were significant for the IG context on all measures (gain scores: 23.9% to 15.7%; supplementary post-tests: 58.6% to 52.5% ; PFNET correlations: .583 to .454; all p's $< .05$), the PS context showed no significant or marginal declines from initial to delayed post-test scores (gain scores: 16.5% to 11.1%; supplementary post-tests: 57.3% to 54.4%; PFNET correlations: .510 to .485; all *n.s.;* see Table 1). (Slight differences in means between these initial post-test scores and those previously reported are because one subject did not participate in the delayed phase). Further, the information gathering context yielded a significantly larger PFNET correlation decay, compared to that yielded by the PS context (IG = .13; PS = .03; $F_{(1, 40)} = 5.002$, $p < .05$).

A significant finding that actually suggests superior performance under the problem solving context upon delayed testing is that a greater number of the expert PFNETs correlated more highly with the PFNETs yielded by the PS context than the PFNETs arising from the IG context (Wilcoxon Signed Ranks Test: $Z = 2.090$, $p < .05$). This is essentially directly opposite to the finding after the initial post-test.

Further findings from the delayed post-test with regard to the effects of pair membership showed that, as in the initial post-test, undergraduates who were paired with graduate students performed significantly better on the ultimate gain score than those in undergraduate-undergraduate pairings (27% vs. 10%; $p < .01$; see Table 2). Indeed, graduate students also performed significantly better on this (delayed) gain score when paired with undergraduates rather than graduate students (18.8% vs. 9.1%; $p = .05$) – and in further contrast to those paired with fellow graduate students, showed no significant decline in scores upon delayed testing. As in the initial phase, non-parametric tests comparing means of all measures indicated that both undergraduate and graduate students performed better in mixed academic pairs than their respective peers in homogenous pairs (Wilcoxon Signed Ranks Test: $Z = 2.201$, $p < .05$).

	Measure	Homogenous Academic Pairs	Mixed Academic Pairs	F	p value
Undergraduates	Gain-Score (from Pre-Test)	0.103	0.273	9.218	.009
	Supplementary Post-Test Score	0.495	0.499	.003	.954
	PFNET Correlation	0.377	0.454	.750	.401
Graduates	Gain-Score (from Pre-test)	0.091	0.188	4.276	.050
	Supplementary Post-Test Score	0.535	0.631	2.277	.144
	PFNET Correlation	0.499	0.544	.357	.556

Table 2: ANOVA of subjects' delayed performance based on academic pair membership.

Gender pairings yielded results similar to those from the initial phase of the study. Non-parametric tests once again revealed that both males and females from same-gender pairs performed better (when tested individually) when one compares all outcomes. Analysis of variance also revealed that males performed significantly better on delayed gain scores when their learning took place in same-gender, rather than mixed-gender, pairs (19.0% vs. 9.4%; $p < .05$; see Table 3).

To assess near transfer in the delayed phase of the study, subjects were given two new problems to solve that were similar to the clinical problems in the problem-solving context of the initial phase. No context-based difference for either exercise could be demonstrated by either analysis of variance or paired sample t-tests between subjects' performances based on the context in which the material was originally learned (means: PS = 69.3% v. IG = 67.2%, n.s.). However, a significant correlation was observed between subjects' performance on the two problem cases ® = .592, $p = .005$).

	Measure	Homogenous Gender Pairs	Mixed Gender Pairs	F	p value
Females	Gain-Score (from Pre-Test)	0.142	0.098	.580	.456
	Supplementary Post-Test Score	0.545	0.533	.038	.848
	PFNET Correlation	0.567	0.494	1.619	.220
Males	Gain-Score (from Pre-Test)	0.190	0.094	4.435	.048
	Supplementary Post-Test Score	0.544	0.512	.276	.605
	PFNET Correlation	0.440	0.384	.496	.489

Table 3: ANOVA of subjects' delayed performance based on gender pair membership.

Discussion

This study appears to support earlier research indicating that problem solving, although contributing less to learning in the short term, has the benefit of considerable retention after a period of time. Advantages exhibited due to the information gathering context in the short term disappeared on subsequent testing six months later. These findings, which appear to favor the problem-solving context for concept retention, seem particularly impressive given that the study was conducted as part of a regular college course. After subjects were introduced to the concepts via the experimental contexts, they encountered them again in lectures, laboratory procedures, examination review sessions and (presumably) individual study. All of these variables would be expected to introduce noise and mute the contexts' effects on an individual's retention of the concepts. The demonstration of a difference on overall retention, after the introduction of so many modulating factors, underscores the importance of the conditions under which subjects first encounter new material.

If we extrapolate the delayed post-test data to even longer delays, these results serve to (a) support the claims by advocates of PBL regarding its characteristic of increased retention, and (b) indicate that PFNET correlations may be more sensitive indicators of retention than conventional objective post-tests. Both the fact that PFNET decay from the IG context was greater than that from the PS context, and that the expert PFNET correlations were higher for those from the delayed PS context, than from the IG context, support the use of this assessment as a learning outcome.

The claim that problem-solving as a method of instruction leads to a greater ability to apply concepts for future problems did not find much support in this study. However, this effect may be more difficult to demonstrate when subjects are exposed to only one example of a clinical problem, rather than several, as is commonly practiced in problem-based learning. Furthermore, it is conceivable that the additional instructional exposure to the concepts introduced in the experiment may have led to a greater ability to apply these concepts in later encounters with clinical problems. These possibilities represent an area for further study, of course.

In summary, it appears that information gathering contributed to greater initial learning as assessed by this study's outcome measures. Further, it appears that problem-solving leads to somewhat greater retention of what was initially learned. It must be pointed out, though, that although there was less decay in the PS context performance, it still never achieved even marginally significant superiority over the delayed performance of the IG context, so further research is necessary to determine if the information gathering context's measures would continue to decline at a higher rate than those yielded by the problem solving context. Finally, advantages for paired subject performance based on academic heterogeneity and gender homogeneity were stable over both post-testing and delayed post-testing.

References

Albanese, M. A. & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, 68, 52-81.

Anderson, J. (1987). Skill Acquisition: Compilation of Weak-Method Problem Solutions. *Psychological Review*, 94, 192-210.

Coppa, G. & Tancred, E. (1995). *BrainStorm: An Interactive Neuroanatomy Atlas* [Computer Program]. Developed by SUMMIT Technology Group at Stanford University. St. Louis, MO: Mosby .

Durso, F. T., Rea, C. B., & Dayton, T. (1994). Graph-theoretic confirmation of restructuring during insight. *Psychological Science*, 5, 94-98.

Eisenstaedt, R. S., Barry, W. E., & Glanz, K.(1990). Problem-based learning: Cognitive retention and cohort traits of randomly selected participants and decliners. *Academic Medicine*, 65, 511-512.

Goldsmith, T. E. & Johnson, P. J. (1990). A structural assessment of classroom learning. In Roger W. Schvaneveldt, (Ed.), *Pathfinder Associative Networks: Studies in knowledge organization* (pp. 241-254). Norwood, NJ: Ablex Publishing Corp.

Norman G. R., & Schmidt, H. G. (1992). The psychological basis of problem-based learning: a review of the evidence. *Academic Medicine*, 67, 557-65.

Schvaneveldt, R. W., Editor (1990). *Pathfinder Associative Networks: Studies in Knowledge Organization*. Ablex Publishing Corp; Norwood, NJ. Series title: Ablex series in computational sciences.

Underwood, G., Jindal, N., & Underwood, J. (1994). Gender differences and effects of co-operation in a computer-based language task. *Educational Research*, 36, 63-74.

Weidner, J., Ranney, M., & Diamond, M. (1999). Knowledge organizations resulting from pairs' problem-solving versus information gathering activities. In B. Collis & R. Oliver (Eds.) *Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 593-597). Charlottesville, VA: AACE.

White, B.Y., Frederiksen, J.R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16, 3-118.