Educational Psychology, Vol. 21, No. 3, 2001

Reciprocal Peer Tutoring: re-examining the value of a co-operative learning technique to college students and instructors

KENT A. RITTSCHOF & BRYAN W. GRIFFIN, *Georgia Southern University, USA*

ABSTRACT *To examine co-operative learning between pairs of college students in the field of education, the reciprocal peer tutoring (RPT) procedure was used in two experiments with 97 graduates (Experiment 1) and 100 undergraduates (Experiment 2). Contrary to initial studies using a different population of college students, results in both experiments indicated that RPT failed to improve students' understandings of course material compared to an individualised study task. In addition, RPT neither increased students' feelings of self-efficacy nor did it decrease students' levels of test anxiety relative to the control condition. However, participants overwhelmingly reported RPT to be helpful for studying the course content. Results suggest careful consideration of ecological validity in research and expected gains in practice.*

Research demonstrates peer tutoring, a type of co-operative learning, to be an effective strategy for increasing student achievement at various educational levels (Bargh & Schul, 1980; Jenkins & Jenkins, 1985; Greenwood *et al.*, 1988; Sherman, 1991; Slavin, 1991, 1996). Typically, achievement gains result for both students in the peer tutoring partnership—the tutor and the tutee. Interestingly, several studies have found that tutors benefited more than tutees, possibly because of the necessary preparation for the tutoring partnership required of the tutor (Allen & Feldman, 1973; Bargh & Schul, 1980; Annis, 1983; Benware & Deci, 1984). A peer tutoring technique that takes advantage of this tutoring benefit is known as reciprocal peer tutoring (RPT). As developed by John Fantuzzo and his colleagues (Pigott *et al.*, 1986; Wolfe *et al.*, 1986), reciprocal peer tutoring enables each student to assume both roles of tutor and tutee, thus allowing each student to derive the benefits from preparing to teach another student. Students in an RPT dyad must provide instruction, evaluation and reinforcement to one another, thereby encouraging mutual assistance and social support for each other (Pigott *et al.*, 1986; Fantuzzo *et al.*, 1989b; Fantuzzo *et al.*, 1992). Thus, RPT

ISSN 0144-3410 print; ISSN 1469-5820 online/01/030313-19 © 2001 Taylor & Francis Ltd
DOI: 10.1080/01443410120065504

appears to be a potentially useful technique for instructors to consider as part of their course plans. Prior to implementing the technique, however, many college instructors may want to know whether such a cooperative structure is clearly more beneficial than an individualised structure. That is, do the potential benefits of peer assistance and social support, while learning course material outweigh the potential benefits of a more independent focus on course material? The present study addresses the preceding question by examining RPT usage in both undergraduate and gradual level college classrooms.

Co-operative Learning: Theoretical Bases for Success

In examining RPT's effectiveness educators should first consider the theoretical rationale that supports successful forms of co-operative learning. Robert Slavin (1996) described four general theoretical perspectives that explain the beneficial effects of cooperative learning on performance. One perspective involves *motivation*, whereby reward structures allow students to attain their goals if the entire group is successful. A second is the *social cohesion* perspective, which ties the achievement successes to the cohesiveness of the group. This cohesiveness exists when group members care about the success of each of its members. The third and fourth perspectives are the cognitive perspectives, which suggest that interactions in cooperative learning will be beneficial because of the mental processing that occurs as a result. Specifically, the *cognitive-developmental* perspective suggests that higher mental functioning originates in social interaction among students (Vygotsky, 1978) and that working within groups can create cognitive conflict and lead students to disequilibrate immature conceptualisations (Piaget, 1926). Alternately, the *cognitive elaboration* perspective suggests that in co-operative learning, elaboration occurs when one member is explaining to other members. These four theoretical perspectives need not be considered mutually exclusive. In fact, Slavin considers these four perspectives to be complementary.

In his review of co-operative learning research, Slavin (1996) found that providing group rewards and holding students individually accountable for learning are key influences on cooperative learning performance successes. During RPT, group rewards can be both extrinsic, as in course credit or points, and intrinsic, as in feelings of achievement or cohesiveness through working together effectively. However, receiving a reward such as course credit need not be contingent upon the level of performance achieved, while the intrinsic feeling of achievement could vary with the degree of relative success. Hence, whether an additional performance contingent reward is crucial to RPT success is open to question. Slavin, however, further indicated that, while adding group rewards increases the effects of RPT, structured didactic tasks such as RPT may not require performance-contingent group rewards.

Slavin's (1996) other crucial variable, students' individual accountability for learning, is somewhat inherent in the RPT procedure because the purpose is ultimately to prepare each student to individually succeed on an upcoming achievement test. This individual accountability may affect each student's perceived responsibility to the other member. That is, each member of the dyad relies on the other to engage in RPT. Abrami & Chambers (1996) discussed this perception of partner reliance as *role inter-dependence* and described it as being a separate form of motivation from individual accountability. Role inter-dependence may be particularly crucial to RPT success because unlike other forms of co-operative learning where four or five students can make up a group, RPT cannot successfully occur if one individual does not actively

participate. Thus, individual accountability can, in effect, be supported by role inter-dependence and possibly by feelings of group cohesiveness in RPT.

Both group cohesiveness and role inter-dependence are examples of some of the many possible effects of implementing RPT that transcend academic achievement. Other logical examples of such motivational and emotional variables that are important to instructors and students alike are self-efficacy and test anxiety. Self-efficacy involves students' beliefs about their effectiveness in a particular domain or situation (Bandura, 1986). One influence on students' self-efficacy that has particular relevance to RPT is the past successes or failures in the particular domain of interest. So, the question answering and subsequent discussion between peers in RPT should have some influence on students' sense of self-efficacy for the test material. For example, if RPT leads to mastery of course concepts, then it follows that increases in self-efficacy should occur. In fact, research has supported the notion that self-efficacy is promoted when one understands and applies a strategy that enhances achievement (Licht & Kistner, 1986; Schunk, 1989). In addition, co-operative learning strategies are known to promote a greater sense of personal efficacy than individualistic learning experiences (Johnson *et al.*, 1978, 1985).

In contrast to self-efficacy is the debilitating experience of anxiety that students frequently have when confronted with a test of their knowledge. Such anxiety is most pronounced when students' beliefs about competence or self-efficacy are low (Pintrich & DeGroot, 1990). Thus, an understanding of students' levels of test anxiety in conjunction with their levels of self-efficacy can provide a comprehensive indication of their motivational and emotional states in testing situations. These states can then be related to achievement with RPT and viewed as important characteristics of the RPT experience.

Another important characteristic of RPT that makes it an attractive learning method is the student-generated questions about the material to be learned. Question generation is an effective comprehension-fostering, self-regulatory cognitive strategy (Palinscar & Brown, 1984). Thus, RPT can be helpful to students for the co-operative learning benefits, as well as the additional metacognitive benefits inherent in question generation that occur prior to the co-operative interaction.

RPT Effectiveness with College Students

In initial experiments with RPT with college-level participants (Fantuzzo *et al.*, Fox, 1989a,b; Riggio *et al.*, 1991), psychology students were paired with the same partner throughout the semester. Each member of the dyad generated a series of test questions, administered the questions to his or her partner, and provided tutelage as necessary prior to unit examinations. These experiments provided evidence for the positive effects of RPT upon achievement. On average, the observed effect of RPT has been calculated to be about $d = 0.90$ (Fantuzzo *et al.*, 1989a,b; Riggio *et al.*, 1991), where d is the measure of effect size described in Glass *et al.* (1981), and Hunter & Schmidt (1989). An effect size of 0.90 is typically considered to be large (Cohen, 1988).

In a more recent study of RPT with college students, Griffin & Griffin (1997) found that RPT had little to no impact on graduate students' understanding of educational research concepts, compared with an individualised control condition. In a similar study, Griffin & Griffin (1998) also found that RPT did not statistically increase achievement for undergraduate educational psychology students. Perhaps the different levels of quality control used, while implementing RPT can explain this discrepancy

between the two sets of studies. In each of Fantuzzo's experiments, RPT was implemented in the following fashion:

- students were paired with a partner;
- each student developed multiple-choice tests;
- students administered their tests to their partners prior to in-class unit exams, then provided tutelage to their partners on those topics that were problematical.

In the Griffin & Griffin studies, RPT was implemented with some modifications to the procedure as outlined in Fantuzzo *et al.* (1989a,b) and Riggio *et al.* (1991). First, Griffin & Griffin attempted to control for the quality of items written by students in the RPT experimental condition. These students submitted items for instructor review prior to administering the items to their RPT partners. The instructor read each item to ensure that it was not ambiguous and contained only one correct response. Secondly, students were required to administer their tests and tutor each other in-class for 30–45 minutes prior to course examinations. The rationale for these two changes was to increase control over the experimental manipulation. These two alterations differed from Fantuzzo's implementation of RPT because Fantuzzo and associates did not review student developed items prior to their use, and told students to participate in RPT outside of class for an unspecified amount of time.

One could argue that by implementing these modifications, the Griffin & Griffin studies undermined some of the benefits of RPT identified in earlier studies (Fantuzzo 1989a,b; Riggio *et al.*, 1991). For example, because the instructor examined each RPT item prior to its use in the Griffin & Griffin studies, it is likely that the number of invalid items were minimised. Students in the Fantuzzo experiments, however, may have used such invalid items. It is possible that such items could have created important and useful dialog which subsequently resulted in improved understanding of course concepts by RPT partners.

Furthermore, the attempt by Griffin & Griffin to ensure the completion of RPT by requiring students to participate in the test-taking and tutoring components of RPT in-class may have actually resulted in less time spent on RPT. If students completed RPT procedures outside of class, they may have taken more time to probe their partners' understanding of course content and identify weaknesses in each others' understandings. In fact, the hypothesis that total learning time is an important variable to consider in student learning is one that has long been supported by empirical evidence (Bugelski, 1962). Additionally, John Carroll's (1989) model of School Learning, which expresses the degree of student learning as a ratio of time-spent to time-needed, is of particular relevance to RPT. Carroll's model highlights time dependent variables that account for different levels of achievement among students. Two of Carroll's variables are (a) the quality of instruction, which reduces the amount of time-needed to learn the material, and (b) the opportunity to learn, which is the amount of time-allowed for learning. Thus, the control of RPT activities including perceived learning time in the Griffin and Griffin studies may not have encouraged an optimal ratio of learning time-spent to time-needed for the given instructional quality.

A clear understanding of whether implementing RPT during class time and monitoring the procedure led to limitations of RPT effectiveness is important for potential application of the RPT technique in classrooms. It is likely that many instructors who may want to use RPT with students will have them complete their tutoring in class where students are aware that time is limited. Additionally, instructors may be concerned about whether students are actually writing legitimate questions or just going

through the motions so as to appear on-task. For these reasons, closely examining the discrepancy between the two sets of studies is of great practical concern to educators.

Other differences between the experiments of Fantuzzo and those of Griffin and Griffin were participants' ages, educational levels and cultural environments. Students in Fantuzzo's experiments (Fantuzzo *et al.*, 1989a,b; Riggio *et al.*, 1991) were undergraduate psychology students in the far western United States, while graduate students in educational research who were living in the southeastern United States were the primary participants in Griffin & Griffin (1997). Perhaps these differences may also have resulted in the different effects observed. Students enrolled in graduate-level programmes are likely to be more experienced learners and thus, may be more prone to use effective learning strategies on their own than are undergraduate students. Greater learning effectiveness by graduate students will likely reduce the relative benefits of RPT because graduate students' need for an organised procedure for clarification and practice may be less than that of undergraduates. Still, undergraduate education students were used in one experiment of Griffin & Griffin (1998) wherein results corroborated those from their previous study using graduate students. Additionally, the possible motivational differences between education students and psychology students should be considered when interpreting the differences in the outcomes.

Rationale for the Present Study

The purpose of the following experiments was to examine whether the benefits of a co-operative learning procedure used in college-level education courses in the southeastern United States is likely to lead to superior understandings of the course material and better psychological adjustment than the benefits of an independent procedure. We chose to examine the RPT technique because of its apparent suitability to college-level learners and material. In addition, previous studies of RPT have limitations that prevent a satisfactory understanding of the potential benefits and the generalisability of this co-operative procedure.

Thus, using education students as subjects, the following experiments tested whether the dramatic discrepancies in RPT effects with college students, observed between the two sets of experiments reported above, resulted from the in-class modifications to Fantuzzo's RPT procedures (Fantuzzo *et al.*, 1989a,b; Riggio *et al.*, 1991) used in the Griffin & Griffin (1997, 1998) studies. In addition, we chose to use a completely randomised experimental design, similar to Fantuzzo's studies, rather than the quasi-experimental design used predominantly in the Griffin & Griffin studies. We also used an individualised control condition that allows for a high degree of generalisability because of its ecological validity and comprehensive nature. We were interested in learning whether these conditions would result in differences in academic achievement, test anxiety and academic self-efficacy. Our interest in these possible differences lies in both the theoretical aspects of cooperative learning and in the practical considerations of implementing RPT in college classrooms under typical classroom constraints. For instance, we are interested in whether implementing RPT is worth the time, effort and the other resources needed.

Given Fantuzzo and associates' positive findings and the rationale for RPT effectiveness stated above, we anticipated that both Fantuzzo's out-of-class version of RPT and Griffin & Griffin's in-class version of RPT would result in higher average scores for achievement and self-efficacy, and lower average scores for test anxiety than the control group. We also predicted that Fantuzzo's version would result in higher average scores

for achievement and self-efficacy, and lower scores for test anxiety than Griffin & Griffin's version. That is, we predicted that implementing RPT would be valuable to the college students and instructors, particularly under Fantuzzo's implementation conditions.

Experiment 1

Method

Participants. Students enrolled in five sections of a graduate-level introductory educational research course offered at a medium-sized, regional university in the southeastern United States participated in the experiment. A total of 97 students participated. The average student age was about 35, approximately 90% were Caucasian and 75% were female. The second author taught all sections of the educational research course.

Instruments. A 45-item test was developed to serve as the posttest measure of achievement. Content validity was established by ensuring that each item included in the test clearly matched a specified performance objective for the course. A 24-item pretest was also developed, using items from the post-test and was used to measure students' initial knowledge of course content.

Test anxiety and academic self-efficacy were measured using the sub-scales presented by Pintrich & De Groot (1990). These scales provide a general, domain-level measure (Bandura, 1997) of academic self-efficacy and test anxiety. For example, statements such as 'When I take a test I think about how poorly I am doing compared with other students', and 'I feel my heart beating fast when I take an exam' are used to assess domain-level test anxiety, and statements like 'I'm certain I can master the skills being taught in this class' and 'I'm confident I can do an excellent job on the assignments and tests in this course' measure domain-level academic self-efficacy. Students provide responses on a seven-point Likert scale where 1 indicates the statement is 'Not at all true of me', and 7 the statement is 'Very true of me'. Pintrich & De Groot provided strong evidence of reliability and construct validity for this instrument. Further evidence can be found in Pintrich *et al.* (1991). The data collected in this study also revealed high levels of internal consistency ($\alpha = 0.94$ for test anxiety and $\alpha = 0.94$ for academic self-efficacy). The correlation between test anxiety and academic self-efficacy for the data collected in these experiments was -0.36.

Test anxiety and academic self-efficacy were also measured by the Self-efficacy and Test Anxiety Scale (STAS; Griffin, 1994), which provides measures of situation-specific test anxiety and self-efficacy (Spielberger & Vagg, 1995). STAS contains 17 statements to which students respond on a seven-point Likert scale like that found with Pintrich & De Groot's (1990) scales. Nine of the statements measure situational, test-specific test anxiety (e.g. 'Thinking about this exam makes my stomach feel tight, upset, or in knots'; 'Thinking about how I might do on this exam makes me nervous'), and eight statements measure situational, test-specific self-efficacy (e.g. 'I believe I will do well on the upcoming exam'; 'I am sure I can answer some of the more challenging or difficult questions on this exam'). Internal consistency for the data collected in these experiments was high: 0.96 for test anxiety and 0.93 for self-efficacy. Griffin (1994) reported that the instrument had evidence of factorial and construct validity. When a factor analysis was performed for the data collected in this current study, all items loaded distinctly and appropriately on one of the two constructs. Furthermore, scores

for test-anxiety and self-efficacy correlated negatively at -0.52, which corresponds with previous research on the relationship between these two variables (Pintrich *et al.*, 1991). In addition, the correlation between STAS's test anxiety and Pintrich & De Groot's measure of test anxiety was 0.79, and the correlation between STAS self-efficacy and Pintrich & De Groot's measure of self-efficacy was 0.61. Finally, the scores provided by STAS provided slightly stronger correlations with the posttest than did scores from Pintrich & De Groot's sub-scales. The correlations with the posttest were -0.27 with the STAS test anxiety and -0.15 with Pintrich & De Groot's test anxiety sub-scale, and 0.28 with the STAS self-efficacy, and 0.26 for Pintrich & De Groot's self-efficacy sub-scale.

A post-experimental questionnaire was developed to gather reflective-self reports of participants' thoughts regarding the RPT experience. There were four items on the questionnaire as follows:

- Do you think you benefited from writing the RPT questions? If so, how do you think writing the RPT questions helped you in this class?
- Do you think taking your partner's tests and tutoring your partner helped you in this class? If so, how did these activities help you?
- Did you find talking with your partner helpful or distracting?
- Do you think a procedure like this (writing short tests, administering and taking tests, and tutoring) would be helpful in your other classes?

Questions were printed two to a page on two sides of an $8.5' \times 11'$ sheet of white paper with blank spaces between questions for responses to be written.

Design and Procedures. A pre-test–post-test control group design (Campbell & Stanley, 1966) was utilised. Students were each assigned randomly to one of three groups. Analysis of the pre-measures (achievement, test anxiety and academic self-efficacy) and GPA showed that the randomly formed groups did not differ statistically on any of the measures. Descriptive statistics for the pre-measures for Experiment 1 are reported in Table I. The treatment and control groups were:

Out-of-class RPT. Students participating in this version performed RPT in a manner that closely matched the RPT procedures specified by Fantuzzo and colleagues. Specifically, students were:

TABLE I. Descriptive statistics for initial measures in Experiment 1 (graduates)

	Achievement pretest	Academic self-efficacy	Test anxiety	UGPA
In-class RPT	M = 33.58	M = 5.72	M = 3.87	M = 3.11
($n = 34$)	SD = 8.82	SD = 1.09	SD = 1.74	SD = 0.41
Out-of-class RPT	M = 35.55	M = 5.49	M = 3.98	M = 3.19
($n = 32$)	SD = 8.60	SD = 0.87	SD = 1.38	SD = 0.32
Control	M = 35.55	M = 5.43	M = 4.15	M = 3.18
($n = 31$)	SD = 9.73	SD = 0.99	SD = 1.20	SD = 0.44
$F(2,94)$	0.47	0.80	0.30	0.34

Note: None of the groups were statistically different, at either the 0.05 or 0.10 level of significance, on any of the pre-measures. Pretest scores in percent correct. Self-efficacy and test-anxiety were measured from Pintrich & De Groot's (1990) instrument.

- randomly assigned a partner at the beginning of the quarter;
- directed to develop a 10-item multiple-choice test for and prior to each in-class course examination, and to develop a separate answer sheet that explained the correct response to each item;
- instructed to administer their 10-item test to their partner sometime prior to the course examination and provide tutelage as needed;
- submit their 10-item test and answer sheet on the day of the examination.

This group of students was not monitored by the instructor during the tutoring process, but their materials were examined to insure that students completed the tasks. On examination days, these students waited in a separate area if they arrived early, until the in-class RPT students were finished and ready to begin the test.

In-class RPT. Griffin & Griffin (1997, 1998) altered RPT as it was originally designed by Fantuzzo and colleagues. Their modified version incorporated the following steps:

- students were randomly assigned a partner at the beginning of the quarter;
- students were instructed to write 10 multiple-choice test items (with explanations for correct responses) for each in-class course examination;
- students then submitted all items to the instructor for review approximately 1 week prior to the unit examination (the instructor read each item to ensure it had a correct response and was not ambiguous);
- students completed RPT activities (i.e. test administration and tutoring) during class time on exam day, immediately before unit examinations;
- upon completion of RPT prior to an examination, students submitted their tests and written explanations of correct responses to the instructor.

Students were allowed as much time as they felt necessary to complete the procedure, but none required more than 45 minutes. Students normally required between 25 and 40 minutes to complete RPT. The instructor monitored each in-class group to insure individual accountability during the tutoring session. Essentially, the In-Class RPT differed from the Out-of-Class RPT in that it required students to submit RPT items for instructor review prior to using them, students administered their tests and tutored each other during class time before each course examination, and the instructor monitored the tutoring progress.

Control Group. These students were each required to keep a journal of their study activities and reflections relative to the course content. This control condition was designed to represent a useful individualised activity that related to course content and that did not involve co-operative interaction among students. Students were asked to submit a copy of their journal prior to each unit examination. Journal submissions were examined to insure that students completed the task. On exam days, these students waited in a separate area if they arrived early, until the in-class RPT students were finished and ready to begin the exam.

On the first day of class the achievement pretest was administered to all students. Also, Pintrich & De Groot's (1990) test anxiety and academic self-efficacy sub-scales were administered on the first night of class to establish a baseline measure of domain-level test anxiety and academic self-efficacy, and were re-administered on the last night of class. The initial scores obtained from the first administration were used as covariates for the analysis of post-measures of test anxiety and self-efficacy. STAS

TABLE II. Descriptive statistics for initial measures in Experiment 2 (undergraduates)

	Achievement pretest	Academic self-efficacy	Test anxiety	UGPA
In-class RPT	M = 51.98	M = 5.89	M = 4.26	M = 2.78
(n = 30)	SD = 7.18	SD = 0.74	SD = 1.30	SD = 0.44
Out-of-class RPT	M = 52.59	M = 5.54	M = 4.41	M = 2.67
(n = 35)	SD = 6.54	SD = 1.03	SD = 1.50	SD = 0.51
Control	M = 50.71	M = 5.64	M = 3.90	M = 2.91
(n = 35)	SD = 7.47	SD = 0.94	SD = 1.41	SD = 0.50
F(2,97)	0.63	1.27	1.18	2.13

Note: None of the groups were statistically different, at either the 0.05 or 0.10 level of significance, on any of the pre-measures. Pretest scores in percent correct. Self-efficacy and test-anxiety were measured from Pintrich & De Groot's (1990) instrument.

was also administered on the last night of class immediately prior to the post-test. Students participated in RPT three times to match the number of examinations administered in the class.

Experiment 2

Method

Participants. Participants in Experiment 2 represented undergraduate education students enrolled in four sections of a human development course offered at the same university specified in Experiment 1. The sample size for this experiment was 100. The average age of these participants was about 20, approximately 70% were White (approximately 30% were Black) and 84% were female. The first author taught each section of the human development course.

The instruments, procedures and design outlined above for Experiment 1 were replicated in Experiment 2, except that the number of items on both the pre- and post-test was 32 and that students participated in RPT four times. As with Experiment 1, the three groups were not statistically different on any of the premeasures. Descriptive statistics for the pre-measures for Experiment 2 are reported in Table II.

Results

Descriptive statistics and ANCOVA results for each of the five dependent measures are presented in Tables III and IV for Experiments 1 and 2, respectively. For the graduate students in Experiment 1, the three groups did not differ significantly on any of the outcomes. Furthermore, there appears to be no discernable pattern among the means. Neither of the experimental groups obtained means that were consistently better than the control group across the five dependent measures.

For the undergraduates in Experiment 2, there were likewise no significant differences among the post-measures except for situational, test-specific self-efficacy. There was a statistically significant interaction between treatment and the covariate—initial (pre)measure of academic self-efficacy. Even with the interaction effect, the control group had slightly higher levels of self-efficacy than either of the two treatment groups. In general, the interaction appears to reveal little useful information and seems to be an

TABLE III. ANCOV As and summary statistics for dependent measures in Experiment 1 (graduates)

		F				
Source	*d.f.*	Post-test[a]	Situational SE	Situational TA	Domain SE	Domain TA
Treatment (T)	2	0.88	0.45	1.37	0.20	0.63
Pretest (P)	1	3.71	15.89*	84.90*	25.99*	160.49*
T × P	2	2.95	0.34	1.08	0.14	0.81
Residual	91	(99.75)[b]	(1.26)	(2.07)	(1.12)	(1.00)
Out-of-class RPT		$M = 74.67$	$M = 5.13$	$M = 4.28$	$M = 5.53$	$M = 3.75$
($n = 34$)		SD = 9.16	SD = 1.17	SD = 2.09	SD = 1.19	SD = 1.85
		M_{adj}[c] = 74.90	$M_{adj} = 5.05$	$M_{adj} = 4.41$	$M_{adj} = 5.43$	$M_{adj} = 3.87$
Effect size, *d*		− 0.06	0.18	0.08	0.19	0.12
In-class RPT		$M = 78.27$	$M = 5.07$	$M = 4.10$	$M = 5.51$	$M = 3.84$
($n = 32$)		SD = 12.45	SD = 1.19	SD = 2.05	SD = 1.06	SD = 1.77
		$M_{adj} = 78.10$	$M_{adj} = 5.10$	$M_{adj} = 4.12$	$M_{adj} = 5.55$	$M_{adj} = 3.85$
Effect size, *d*		0.25	0.23	− 0.07	0.29	0.11
Control		$M = 75.66$	$M = 4.77$	$M = 4.41$	$M = 5.13$	$M = 3.81$
($n = 31$)		SD = 9.09	SD = 1.24	SD = 1.90	SD = 1.33	SD = 1.45
		$M_{adj} = 75.50$	$M_{adj} = 4.83$	$M_{adj} = 4.26$	$M_{adj} = 5.20$	$M_{adj} = 3.67$

Note: Pretest (P) refers to the appropriate pretest for each dependent measure. Situational SE = situational self-efficacy; Situational TA = situational test anxiety; Domain SE = domain-level self-efficacy; Domain TA = domain-level test anxiety.

[a]Scores in per cent correct.

[b]Mean squared errors.

[c]Adjusted means.

*$P < 0.05$.

anomaly given that of the ten ANCOVAs performed, only one indicated differences existed.

Post-experimental Questions. Following an initial survey of all the responses to each post-experimental question by the two authors, a set of response categories were agreed upon to enable coding of the participants' responses. These categories were based upon theory and research on self-regulated learning (Winne, 1995; Zimmerman & Martinez-Pons, 1986, 1988; Zimmerman & Risemberg, 1997) to help provide a meaningful, well-documented framework for the interpretation of the responses. Self-regulated learning categories that were used to classify the responses included application of learning strategies, self-evaluation, goal-setting, planning, attention-focus and seeking peer assistance (see Appendix). Specific category definitions were adapted from those of Ormrod (1999).

Each response was coded by the two authors, then compared for consistency. Discrepancies were resolved through mutual agreement between the two authors. A random sample of 30 data sets (15 from each experiment) were then rated by a third rater who was unaware of the experimental variables. Multiple categories were allowed in the rating of each response. The inter-rater reliability was examined using two methods in order to examine (a) agreement on some portion of each subject's response and (b) precise agreement on all possible categories reflected in subjects' responses.

TABLE IV. ANCOVAs and summary statistics for dependent measures in Experiment 2 (undergraduates)

		F				
Source	*d.f.*	Post-test[a]	Situational SE	Situational TA	Domain SE	Domain TA
Treatment (T)	2	1.87	6.82*	0.56	1.81	0.84
Pretest (P)	1	27.77*	7.77*	28.86*	11.85*	34.21*
T × P	2	1.39	5.93*	0.69	1.83	0.51
Residual	94	(90.04)[b]	(0.85)	(2.01)	(0.73)	(1.70)
Out-of-class RPT		$M = 75.42$	$M = 4.83$	$M = 3.75$	$M = 5.70$	$M = 3.90$
($n = 30$)		SD = 11.86	SD = 1.19	SD = 1.53	SD = 0.85	SD = 1.52
		M_{adj}[c] = 75.30	M_{adj} = N/A	$M_{adj} = 3.71$	$M_{adj} = 5.62$	$M_{adj} = 3.86$
Effect size, *d*		− 0.45	Varies	0.15	− 0.20	0.16
In-class RPT		$M = 78.66$	$M = 5.04$	$M = 4.12$	$M = 5.67$	$M = 3.62$
($n = 35$)		SD = 10.32	SD = 1.09	SD = 1.62	SD = 1.14	SD = 1.60
		$M_{adj} = 78.10$	M_{adj} = N/A	$M_{adj} = 4.00$	$M_{adj} = 5.73$	$M_{adj} = 3.50$
Effect size, *d*		− 0.19	Varies	0.33	− 0.08	− 0.08
Control		$M = 79.38$	$M = 5.46$	$M = 3.29$	$M = 5.78$	$M = 3.46$
($n = 35$)		SD = 10.00	SD = 0.71	SD = 1.71	SD = 0.72	SD = 1.45
		$M_{adj} = 80.10$	M_{adj} = N/A	$M_{adj} = 3.46$	$M_{adj} = 5.80$	$M_{adj} = 3.62$

Note: Pretest (P) refers to the appropriate pretest for each dependent measure. Situational SE = situational self-efficacy; Situational TA = situational test anxiety; Domain SE = domain-level self-efficacy; Domain TA = domain-level test anxiety.
[a]Scores in per cent correct.
[b]Mean squared error.
[c]Adjusted means.
*$P < 0.05$.

The inter-rater reliability measured as a proportion of categorisation agreement for one or more category per subject was 0.90. The inter-rater reliability measured as the proportion of exact categorization agreement between raters was 0.72. These differences in reliability estimates suggest that the interpretive nature of the reasons provided and the overlap among self-regulatory learning categories should be considered when examining trends illustrated in the percentages of categories represented in participant responses. Table V presents the percentages of categorised student responses among experimental groups in Experiments 1 and 2.

The post-experimental reports suggested an overwhelmingly positive response to RPT and highlighted several categories of reasons supporting students' perspectives. For instance, among the four RPT groups examined in the two experiments, 91.9–100.0% of the respondents reported that writing of the test items was helpful. The predominant reasons given for question writing being helpful were its effects on students' application of learning strategies and their self-evaluation. For undergraduates in Experiment 2, goal setting was also frequently cited. In response to whether test taking and tutoring were helpful, 69.6–94.6% of respondents in both experiments indicated that they were helpful, with undergraduates being somewhat more favourable than graduates. The most consistently provided reasons given by all groups were the effects of seeking peer-assistance, self-evaluation and application of learning strategies. Again, goal setting was also frequently provided as a reason by undergraduates in Experiment 2. In response to whether communication with their

TABLE V. Response percentages for post-experimental questions regarding RPT in Experiments 1 and 2

	Graduate in class[a] ($n = 23$)[b]	Graduate out of class[c] ($n = 28$)	Undergraduate in class ($n = 37$)	Undergraduate out of class ($n = 33$)
Writing items helpful?	95.7 yes	92.9 yes	91.9 yes	100.0 yes
	4.3 no	3.6 no	5.4 no	0.0 no
	0.0 unc.[d]	3.6 unc.	2.7 unc.	0.0 unc.
Reasons it was helpful				
Application of learning strategies	47.4	39.3	52.3	51.9
Self-evaluation	23.7	26.8	9.2	21.2
Goal setting	7.9	1.8	21.5	9.6
Planning	10.5	7.1	3.1	1.9
Attention focus	5.3	7.1	3.1	9.6
Seeking peer assistance	2.6	8.9	4.6	0
Other/does not fit	2.6	3.6	1.5	1.9
No reason given	0	5.4	4.6	3.8
Taking test and rutoring helpful?	69.6 yes	85.7 yes	94.6 yes	90.9 yes
	21.7 no	10.7 no	2.7 no	6.1 no
	8.7 unc.	3.6 unc.	2.7 unc.	3.0 unc.
Reasons it was helpful				
Application of learning strategies	11.4	13.0	19.3	11.3
Self-evaluation	31.4	26.1	21.1	26.4
Goal setting	0	6.5	22.8	18.9
Attention focus	0	2.2	0	7.5
Seeking peer assistance	34.3	37.0	24.6	26.4
Other/does not fit	5.7	6.5	7.0	3.8
No reason given	14.3	8.7	5.3	5.7
Communication with partner helpful?	60.9 yes	75.0 yes	70.3 yes	60.6 yes
	13.0 no	7.1 no	13.5 no	6.1 no
	21.7 unc.	17.8 unc.	16.2 unc.	33.4 unc.
Would RPT be useful in other courses?	65.2 yes	67.9 yes	73.0 yes	72.7 yes
	8.7 no	10.7 no	2.7 no	3.0 no
	26.0 unc.	21.4 unc.	24.3 unc.	24.3 unc.

[a]RPT condition that was closely monitored and conducted during class time.
[b]Cell sizes that differ from total n in each group reflect missing data.
[c]RPT condition that was not closely monitoried and was conducted outside of class.
[d]Responses that reflected uncertainty. Not clearly yes or no.

partner was helpful, 60.6–75.0% reported that it was helpful. Similarly, in response to whether RPT would be useful in other courses, 65.2–73.0% reported that it would be useful, with undergraduates appearing somewhat more positive than graduate students on this issue.

Discussion

Essentially, none of our main predictions involving the performance tests, self-efficacy or test anxiety were supported by either of the two experiments. Noting the strong positive findings reported in Fantuzzo *et al.* (1989a,b) and Riggio *et al.* (1991) for undergraduate students, we anticipated that students participating in Fantuzzo's version of RPT would demonstrate greater achievement, less test anxiety and greater

academic self-efficacy than would the control group. Consistent with Griffin & Griffin (1997, 1998), we found that RPT did not appear to enhance achievement for graduate students of educational research. Also consistent with Griffin & Griffin (1998), we could not replicate the effects for RPT with our undergraduate students.

On the other hand, the post-experimental reports provided a contrasting perspective. Participants overwhelmingly perceived that they benefited from writing RPT questions and that it helped them in the course they were taking. Similarly, they overwhelmingly perceived that taking the tests and tutoring were helpful in the course they were taking and, for most, would likely be helpful in other courses. In contrast to the quantitative findings of this study, we found that the qualitative reports support a number of beneficial effects of the RPT technique for both graduates and undergraduates. It appears that the post-experimental questions were able to tease out potentially valuable perceptions of participants that are not apparent in the quantitative performance outcomes. For example, the strongly positive response regarding the helpfulness of the RPT technique could not have been inferred from the lack of relative benefit indicated by the objective test performance. In addition, the various ways in which participants perceived benefit support the notion that the performance test and the surveys did not reveal some important effects of RPT on the students. Although experimenter effects or social desirability effects (Watkins, 1996) on participant responses are possible, the explanations that participants provided strongly suggest that their responses were based upon substantive reasons. In addition, there were a small number of participants who reported that RPT activities would not be helpful in other classes. These few negative responses among both graduates and undergraduates suggest that there was not an apparent obligation to speak positively of the procedure. Also of interest are the consistent categories of explanations that lend support for some, but not all, theoretical perspectives on cooperative learning.

An examination of four major theoretical frameworks used to explain effects of co-operative learning (Slavin, 1996) shows that participants' reports provide some support for at least two major theoretical frameworks. The cognitive elaboration perspective, which emphasises the elaboration effects that occur when one person is explaining to another or others, is corroborated by the qualitative data. Support for this perspective comes from the consistent reports that were categorised as 'seeking peer assistance'. This category also lends support to a cognitive developmental perspective. In particular, they support the Vygotskian cognitive developmental emphasis on social interaction as a means to effective learning (Vygotsky, 1978). Still, though the participant reports are consistent with these theoretical perspectives, the degree to which students engaged in the reported activities is unclear. For instance, the level at which most RPT students interacted or discussed information may not have been sufficient to yield test performances that were superior to students in the individualised control group.

Although the participant reports did not yield clear support for a reward-based motivational perspective or a social cohesion perspective (Slavin, 1996), other consistent responses revealed several additional areas that highlighted the value students see in the RPT technique. For instance, the RPT effects on self-evaluation, goal setting, planning and attention control are all effects that are of likely interest to instructors considering whether the use of RPT is worthwhile. Likewise, the reports that RPT was not helpful to a few students should also be noted.

The consistent reports of the various self-regulatory learning activities by RPT participants are of particular interest in light of recent investigations into self-regulated

learning (Winne, 1995; Zimmerman, 1995). The participant reports in the present study suggest that forms of self-regulated learning should be systematically examined within the context of RPT. Based upon the present qualitative findings, future investigations should further examine the extent to which self-regulated learning is promoted by RPT and which aspects of self-regulated learning in particular are most influenced. Additionally, based upon the contrast between the quantitative and qualitative findings, investigations could examine how various forms of self-regulated learning affect subsequent use of information learned through RPT.

Considering the contrast between quantitative and qualitative findings, it should also be noted that other co-operative learning research has yielded findings related to this study's outcomes involving performance versus perceptions. Snyder & Sullivan (1995) examined cooperative versus individualised science learning conditions with low, medium and high ability-level seventh graders placed in mixed-ability groups. In their study, all but the high-ability boys preferred the co-operative condition, even though the individualised group performed better than the co-operative group. Thus, even among early adolescents, student perceptions regarding co-operation were not necessarily associated with superior test performance.

RPT Use in College Courses

There are several important ways that the curious findings from the present study can be viewed. This study brings into question the real value of the RPT technique for some college students and instructors. The findings of this study support the possibility that college students' test performance, test anxiety levels and self-efficacy are not necessarily improved in any notable ways by engaging in RPT, as compared with engagement in an individualised study strategy. With our subject sample, we find little reason to agree with the claim of 'a demonstrated ability of the RPT technique to maximize existing classroom resources to enhance academic gains …' (Fantuzzo *et al.*, 1989b, p. 177).

In considering our failure to corroborate some of the beneficial effects of previous studies one might examine the specific task demands of the control groups used in the various studies. For example, while the present study and Fantuzzo *et al.* (1989b) both required students to engage in a writing exercise, the present study required that students write a 'journal' on their reflections and progress with all the material being covered. This type of writing exercise was used to provide an ecologically valid, content-related task that was non-cooperative in nature. Fantuzzo *et al.* (1989b), on the other hand, required their 'independent unstructured' group participants to write about 'a general topic related to the unit exam' (p. 174), rather than on all of the unit topics. Some may believe that their control subjects were consuming study time with an activity that only dealt with a fraction of the material that they would be tested on. It could be argued that this focused writing exercise may have been relatively detrimental to students' test performance and their self-efficacy, while possibly increasing their test anxiety, and that such a control condition may not provide the strongest test of the RPT technique compared with the more exam-related control condition of the present study. Furthermore, the argument could be made that the focused writing control condition could exaggerate the relative effects of RPT and help explain the contrasting lack of effects found when a more comprehensive writing activity was used as a basis for comparison. Similarly, their 'independent structured' control condition (Fantuzzo *et al.*, 1989b, p. 174), which required question generation without test taking, may not be

viewed as a particularly strong test of RPT either. Although independent question generation may be a logical experimental comparison, the activity itself may not be seen as one that encourages a high degree of student reflection on misunderstandings unless combined with training and guidance (Rosenshine *et al.*, 1996).

Another important reason for the differences in findings between studies could be the differences in techniques used to monitor student accountability. In the present study accountability was insured in a manner consistent with typical classroom practices. That is, instructors checked student materials during and after the cooperative work, and they walked around the classroom and insured that students were on-task in the in-class experimental groups. These accountability procedures were considered to be an ecologically valid approach. On the other hand, after the present experiments were conducted John Fantuzzo (personal communication, March 8, 1999) graciously provided some important feedback regarding his studies that was not reported in his publications (Fantuzzo *et al.*, 1989a,b; Riggio *et al.*, 1991). He indicated that accountability was insured through the use of tape recordings of co-operative groups. While we agree that tape-recorded surveillance is an excellent way to insure accountability, we deliberately avoided tape recording cooperative groups because we believe that tape recording:

- prevents meaningful comparison to individualised groups who are not tape recorded (i.e. the RPT treatment and audiotaping are confounded);
- is not ecologically valid in most co-operative learning situations.

Note that based upon this information provided after the present studies were conducted, the Out-of-Class condition that we said closely matched the procedure used by Fantuzzo and his colleagues is different from the procedures that he and his colleagues actually used because of the tape-recording. Our Out-of-Class procedure only matches what was reported in the publications (Fantuzzo *et al.*, 1989a,b; Riggio *et al.*, 1991). Fantuzzo's clarification to us is extremely helpful in providing appropriate context for the interpretations of the various studies that have been conducted on RPT.

Thus, instructors can interpret the findings of RPT studies based upon the likely implementation methods and the likely alternatives to RPT that they may actually consider for their college-level students. While focused writing and question generation may be useful requirements in many courses, we suggest that in practice they are incomplete alternatives to a comprehensive study procedure. However, we also believe that there is now strong evidence that the differences among control groups discussed above are not the most important influences on the outcome discrepancies of the various RPT studies. Considering the control groups used in Griffin & Griffin (1997, 1998) that consisted of either no additional activity or a focused writing activity, and their corresponding lack of achievement gains with RPT, we find that the possible arguments regarding control conditions discussed above are not strongly supported by the available data. Hence, the accountability procedures (e.g. audiotaping or otherwise monitoring), the student population differences and the course content differences among the contrasting studies become more important considerations when interpreting the discrepant outcomes.

A positive perspective on the quantitative findings is that this study supports all the previously known findings that RPT does not appear to be detrimental to undergraduates or graduates in learning course material, as we expected. That is, students using RPT did not perform statistically worse than students engaged in an individualised activity. College instructors and students have reason to be confident that RPT

is a relatively harmless strategy for review and practice. As an alternative to other learning approaches, RPT can be seen as a means to adding variety and a co-operative goal structure to the weekly class activities that college students can engage in. Additionally, we have no support for the notion that RPT during class time will be inferior to RPT outside of class. Under the conditions of the present experiment, the relatively unrestricted time allowed for engaging in RPT outside the classroom did not provide any test performance gains or improvements in anxiety and self-efficacy over the condition that limited the students' perceived time within the classroom. Thus, we have no support for the previously suggested possibility that Fantuzzo's studies yielded greater gains than those of Griffin & Griffin because of an increased opportunity to practice and review with the RPT technique.

Conclusions

Based on the lack of performance gains with RPT found in the present study we suggest that college instructors cautiously weigh the cost-to-benefits ratio when considering the use of cooperative learning strategies such as RPT in their classrooms. Instructors may find the costs of implementing RPT and of utilising valuable time to outweigh the benefits for college students if the reasons for using RPT are primarily based on achievement gains. While we have no basis to dispute the value of co-operative strategies such as RPT for children (Fantuzzo *et al.*, 1995), it is possible that many college-level students have acquired the metacognitive means to allocate their limited time to more test-focused learning than that which is allowed with the RPT technique.

Furthermore, the present findings lead us to suggest that educators and researchers carefully consider the bases for comparison when interpreting achievement gains with co-operative learning techniques. It is likely that many instructors have implemented cooperative learning techniques based upon reports of achievement gains, without having scrutinised the experimental and control groups' activities of studies reporting such gains. We suggest that when RPT is being used primarily for achievement gains, instructors should evaluate the ecological validity of the studies that provide a basis for expected improvements in test scores.

Still, the RPT technique appears to enhance the enjoyment of learning and stimulate various types of self-regulation. Thus, the qualitative reports seem to provide evidence in favour of the claim that RPT can 'increase satisfaction with the learning environment …' (Fantuzzo *et al.*, 1989b, p. 177). Considering that RPT did not degrade objective performance, some instructors may find that the use of RPT provides motivational outcomes and co-operative experience benefits that outweigh the costs. Also, another interpretation of the contrast between test performances and the qualitative reports is that the high level of student satisfaction and perceived benefit results in part from the depth and breadth of understanding that students may be left with after engaging in RPT, regardless of their test performances. That is, RPT may have possibly resulted in student learning of relevant incidental information that went untested. Additionally, the simple fact that many students thought it was helpful may have facilitated incidental learning and increased positive emotions about the course topics.

Based upon the findings presented here, we also suggest that researchers conduct investigations that specifically examine whether various individualised study techniques yield greater effects for college students than does RPT. In addition to objective test performance, test anxiety and self-efficacy, researchers should examine other meaningful measures including longer-term retention, subjective measures of performance,

types of self-regulation and interest in the subject. Such measures may reveal additional benefits of co-operative techniques such as RPT. Clearly, the relationships among student perceptions, motivational outcomes and learning outcomes in co-operative learning environments is worthy of continued investigation.

Correspondence: Kent A. Rittschof, Curriculum, Foundations, and Research, P.O. Box 8144, Georgia Southern University, Statesboro, GA 30460, USA.
(E-mail: gskar@gasou.edu)

REFERENCES

Abrami, P.C., & Chambers, B. (1996). Research on cooperative learning and achievement: comments on Slavin. *Contemporary Educational Psychology, 21*, 70–79.

Allen, V.L., & Feldman, R.S. (1973). Learning through tutoring: low achieving children as tutors. *Journal of Experimental Education, 42*, 1–5.

Annis, L.F. (1983). The processes and effects of peer tutoring. *Human Learning, 2*, 39–47.

Bandura, A. (1997). *Self Efficacy: the exercise of control*. New York: W.H. Freeman and Company.

Bandura, A. (1986). *Social Foundations of Thought and Action: social cognitive theory*. Englewood Cliffs: Prentice Hall.

Bargh, J.A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology, 72*, 593–604.

Benware, C.A., & Deci, E.L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal, 21*, 755–765.

Bugelski, B.R. (1962). Presentation time, total time, and mediation in paired-associate learning. *Journal of Experimental Psychology, 63*, 409–412.

Campbell, D.T., & Stanley, J.C. (1966). *Experimental and Quasi-experimental Designs for Research*. Chicago Rand McNally.

Carroll, J.B. (1989). The Carroll model: a 25 year retrospective and prospective view. *Educational Researcher, 18*, 26–31.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Hillsdale: Erlbaum.

Fantuzzo, J.W., Dimeff, L.A., & Fox, S.L. (1989a). Reciprocal peer tutoring: a multimodal assessment of effectiveness with college students. *Teaching of Psychology, 16*, 133–135.

Fantuzzo, J.W., Riggio, R.E., Connelly, S., & Dimeff, L.A. (1989b). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: a component analysis. *Journal of Educational Psychology, 81*, 173–177.

Fantuzzo, J.W., King, J.A., & Heller, L.R. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: a component analysis. *Journal of Educational Psychology, 84*, 331–339.

Fantuzzo, J.W., Davis, G.Y., & Ginsburg, M.D. (1995). Effects of parent involvement in isolation or in combination with peer tutoring on student self-concept and mathematics achievement. *Journal of Educational Psychology, 87*, 272–281.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in Social Research*. Beverly Hills: Sage).

Greenwood, C.R., Carta, J.L., & Hall, V. (1988). The use of peer tutoring strategies in classroom management and educational instruction. *School Psychology Review, 17*, 258–275.

Griffin, B.W. (1994). *Self-efficacy and test anxiety scales*, Unpublished manuscript, Georgia Southern University.

Griffin, B.W., & Griffin, M.M. (1997). The effects of reciprocal peer tutoring on graduate student achievement, test anxiety, and academic self-efficacy. *Journal of Experimental Education, 65*, 197–209.

Griffin, M.M., & Griffin, B.W. (1998). An investigation of the effects of reciprocal peer tutoring on student achievement, test anxiety, and academic self-efficacy. *Contemporary Educational Psychology, 23*, 298–311.

Hunter, J.E., & Schmidt, F.L. (1989). *Methods of Meta-analysis*. Beverly Hills: Sage.

Jenkins, J., & Jenkins, L. (1985). Peer tutoring in elementary and secondary programs. *Focus on Exceptional Children, 17*, 3–12.

Johnson, D.W., Johnson, R., & Scott, L. (1978). The effects of cooperative and individualized instruction on student attitudes and achievement. *Journal of Social Psychology, 104*, 207–216.

Johnson, D.W., Johnson, R., Pierson, W., & Lyons, V. (1985). Controversy vs. concurrence seeking in multigrade and single-grade learning groups. *Journal of Research in Science Teaching, 22*, 835–848.

Licht, B.G., & Kistner, J.A. (1986). Motivational problems of learning disabled children: individual differences and their implications for treatment. In: J. K. Torgeson, & B. W. L. Wong (Eds) *Psychological and Educational Perspectives on Learning Disabilities*, pp. 225–255. Orlando: Academic Press.

Ormrod, J.E. (1999). *Human Learning*. Upper Saddle River: Prentice-Hall.

Palinscar, A.S., & Brown, A.L. (1984). Reciprocal teaching of comprehension fostering and comprehension-monitoring activities. *Cognition and Instruction, 2*, 117–175.

Piaget, J. (1926). *The Language and Thought of the Child*. New York: Harcourt Brace.

Pigott, H.E., Fantuzzo, J.W., & Clement, P.W. (1986). The effects of reciprocal peer tutoring and group contingencies on the academic performance of elementary school children. *Journal of Applied Behavior Analysis, 19* 93–98.

Pintrich, P.R., & De Groot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40.

Pintrich, P.R., Smith, D.A., Garcia, T., & McKeachie, W.J. (1991). A manual for the use of the motivated strategies for learning questionnaire (MSLQ), *Tech. Rep. No. 91-B-004*. Ann Arbor: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

Riggio, R.E., Fantuzzo, J.W., Connelly, S., & Dimeff, L.A. (1991). Reciprocal peer tutoring: a classroom strategy of promoting academic and social integration in undergraduate students. *Journal of Social Behavior and Personality, 6*, 387–396.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: a review of intervention studies. *Review of Educational Research, 66*, 181–221.

Schunk, D.H. (1989). Self-efficacy and cognitive skill learning. In: C. Ames, & R. Ames (Eds) *Research on Motivation in Education*, vol. 3, pp. 13–44. New York: Academic Press.

Sherman, L.W. (1991). *Cooperative Learning in Post Secondary Education: implications from social psychology for active learning experiences*, paper presented at the annual meeting of the American Educational Research Association, 3–7 April, Chicago, IL.

Slavin, R.E. (1991). Synthesis of research on cooperative learning. *Educational Leadership, 48*, 71–82.

Slavin, R.E. (1996). Research on cooperative learning and achievement: what we know, what we need to know. *Contemporary Educational Psychology, 21*, 43–69.

Snyder, T. & Sullivan, H. (1995). Cooperative and individual learning and misconceptions in science. *Contemporary Educational Psychology, 20*, 230–235.

Spielberger, C.D., & Vagg, P.R. (1995). Test anxiety: a transactional process model. In: C. D. Spielberger & P. R. Vagg (Eds) *Test Anxiety: theory, assessment, and treatment*, pp. 3–14. Washington DC: Taylor & Francis).

Vygotsky, L.S. (1978). *Mind in Society: the development of higher psychological processes*. In: M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds), p. 86. Cambridge: Harvard University Press.

Watkins, D. (1996). The influence of social desirability on learning process questionnaires: a neglected possibility? *Contemporary Educational Psychology, 21*, 80–82.

Winne, P.H. (1995). Self-regulation is ubiquitous, but its forms vary with knowledge. *Educational Psychologist, 30*, 223–228.

Wolfe, J.A., Fantuzzo, J.W., & Wolfe, P.K. (1986). The effects of reciprocal peer management and group contingencies on the arithmetic proficiency of underachieving students. *Behavior Therapy, 17*, 253–265.

Zimmerman, B.J. (1995). Self-regulation involves more than metacognition: a social cognitive perspective. *Educational Psychologist, 30*, 217–221.

Zimmerman, B.J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*, 614–628.

Zimmerman, B.J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology, 80*, 284–290.

Zimmerman, B.J., & Risemberg, R. (1997). Self-regulatory dimensions of academic learning and motivation. In: G. D. Phye (Ed.) *Handbook of Academic Learning: construction of knowledge*. San Diego: Academic Press.

Appendix: Coded Category Definitions and Representative Participant Examples of Actual Responses

Self-evaluation: assessing the progress or outcome of one's work. 'It showed me what I knew or didn't know very well.'

Application of learning strategies: selecting and using appropriate ways of processing material. 'By rehearsing the material over again.'

Goal setting: identifying a desired end result for the learning activity. 'It helped me to identify areas where I might need to study more.'

Planning: determining how best to use the time available for the learning task. 'They helped by making me study more.'

Attention control: maximising attention on the learning task. 'The tests help me focus on important information.'

Seeking peer assistance: soliciting another person's help, teaching, and/or perspectives. 'She would tutor me and I learned from her reasoning.'