A pilot intervention to improve the structural quality of exam essay writing in UK undergraduate psychology students

Vincent Connelly, Oxford Brookes University,
Julie E. Dockrell, University of London &
Jo Barnett, South Bank University

Psychology undergraduates need to produce good quality essays in order to succeed at university. Students find the transition to university writing difficult. Using a rubric, a profile of student weakness in psychology essay writing was described. The students were generally poor at the structural organisation of their essays. A pilot intervention to improve essay structure was designed based on the profile of student weakness. The intervention was trialled with a group of 23 psychology undergraduate students. Those students who took part in the intervention programme had higher exam essay marks with longer and more clearly structured essays after the programme than a matched group of controls. The intervention group students produced essays that linked evidence more clearly to topics and provided better conclusion sections. Psychology tutors were able to show improvement in student writing with a short explicit lesson based on a discipline-specific profile of writing. However, there was no improvement in vocabulary for linking topics and the structure of the introductions remained poor. Ideas for the further development of this intervention are discussed.

THERE HAS been much recent concern about the perceived poor academic writing skills of undergraduate students (Hartley, 2001). Employers rate good writing as one of the most important set of transferable skills developed in higher education, and undergraduates must master the ability to produce a good written composition. However, many students have rated producing writing as one of the most difficult of university assignments (Singleton, 1999). Indeed, research has shown that producing written text requires the coordination of a large number of complex cognitive and metacognitive skills (Hayes, 1996; Hayes & Flower, 1986) and can take many years to develop to maturity. It has been shown that some students are entering higher education with writing skills that may compromise their ability to achieve good grades (Connelly *et al.*, 2005) or with particular difficulties that may affect their production of written text (Connelly *et al.*, 2006).

Recent research into writing has shown that students also need to develop writing skills that are specific to the genres they will encounter at university (Torrance, 1996) and that are specific to the disciplines they study (Monroe, 2002). In psychology, this primarily means learning to write the persuasive essay since the majority of undergraduate students still receive the most significant proportion of their marks from persuasive essays and students expend a lot of time and effort on preparing them (Maclellan, 2001; Norton *et al.*, 1999). Addressing these students' needs entails both an understanding of the writing process and an awareness of the potential difficulties that specific cohorts experience with discipline specific discourses (McClune, 2004).

University teaching staff have traditionally shunned the explicit teaching of writing skills (Channok, 2000). However, recent research has shown that transfer from writ-

© The British Psychological Society 2006

ing strategies appropriate in school education, such as narrative, to the persuasive writing prevalent in higher education is far from automatic (Lavalle *et al.*, 2002). Therefore, there may be a need for universities to tackle this potential stumbling block for students.

There has been a recent move in UK higher education towards including more writing practice in the academic curriculum itself as way of infusing writing into the student workload. This movement has derived from the United States Writing in the Disciplines (WID) initiative. Discipline-specific writing sessions are introduced throughout the undergraduate curriculum. There is evidence that WID can improve student writing through extra practice effects (Ashbaugh *et al.*, 2002; Johnstone *et al.*, 2002; Riordan *et al.*, 2000), but there is less evidence that WID improves student thinking and knowledge construction (Bangert-Drowns *et al.*, 2004). Typically, the implementation of WID is a large-scale undertaking more often suitable to be driven at school or university level rather than at a departmental level (Hennessy & Evans, 2005; Somerville & Creme, 2005). Hence, it is only found in a few UK higher educations institutions at present. None the less the underpinning rationale of improving discipline specific writing can be implemented at any level – course, department, faculty or university. The present study draws on analogous principles to the WID movement in that it was targeted, had a discipline-specific focus and was tailored to the specific needs of the students by being taught within the normal teaching curriculum by discipline-specific staff.

To implement wider changes within a department and its curriculum it is important to demonstrate the efficacy of an intervention. We were set the challenge of establishing an evidence base in a short space of time, with limited resources, that could convince staff of the value of writing in the curriculum. To support our approach we drew on a large body of experimental research showing that explicit and focused direct instruction in discipline-specific writing can produce rapid improvements in student writing that is both effective and long lasting (see Graham, in press, for a review and De La Paz (2005) for a discipline-specific example). This body of research from the United States has focused primarily on schoolchildren and the learning disabled and takes a sociocognitive model of self-regulated learning as the theoretical basis for intervention. The writing interventions have been shown to be effective in older secondary school children (e.g. De La Paz & Graham, 2002) and so seemed to be potentially transferable to undergraduate students since developmental change in writing is very slow compared to other literacy skills (Perkins, 1985).

Discipline-specific instruction in writing would also allow students and tutors to share an understanding about what counts as 'good writing'. Channok (2000) has shown that students frequently misinterpret the feedback comments given to them by tutors on their writing assignments. This showed that students were not sharing the tutors' underlying assumptions about writing. Norton and Norton (2001) have also shown that undergraduate perceptions about how they are achieving departmental essay writing criteria may be very different from tutor perceptions.

A number of research studies have also shown that students who have problems with writing often do not have well-defined goals when they begin a writing task (De La Paz *et al.*, 1998, Nussbaum & Kardash, 2005). They attempt to write a difficult text without being specific about what they aim to write or how they write about it. This is a problem, as it has been found that when the goal of a task is ill defined much of the planning process in writing consists of goal specification (Hayes & Flower, 1980). If students are not clear what tutors understand as 'good writing' then their writing goals will be not be in line with what a tutor considers good writing for their discipline. Therefore, time should be taken to spell out shared goals and assumptions about 'good writing' between student and tutor.

Expert writers devote considerable effort to setting goals in their writing (Bereiter & Scardamalia, 1987) and well-specified writing goals provide clear direction. For example, Hull (1981) set specific text production goals for college students and claimed that they led to increased length of journal entries. These goals have to be specific to the task. Matsuhashi and Gordon (1985) showed that poor college writers set specific revising goals (they had to generate five ideas about the text) improved in their writing more than students given a general revising goal (make sure you improve your paper). This approach has also shown success with learning disabled populations. Ferretti *et al.* (2000) conducted a study where it was demonstrated that providing an appropriate set of elaborated goals (an overall goal comprising a number of subgoals) for learning disabled students led to improvements in writing quality. Nussbaum and Kardash (2005) also carried out a similarly successful intervention on undergraduate text generation by providing appropriate goal instruction. The final advantage of providing specific goals is that the implicit knowledge tutors often take for granted is made explicit to students (Channok, 2000).

We therefore set out to design an intervention based on the principle of providing a specific goals framework to support the structure of student persuasive writing in psychology and to evaluate its efficacy in a cohort of UK second year psychology undergraduate students. Research on children has shown that goal-based interventions can be successful in a short period of time by improving students writing self-regulation, and so have long lasting benefits. The intervention was carried out to raise the profile of writing in the curriculum and to demonstrate the ease of integrating writing skills into the psychology curriculum.

Method

Design

This project was designed to test an intervention to improve undergraduate persuasive writing through providing specific writing goals. The design was a pre–post test with an intervention group and a non-intervention control group. We planned to have a placebo control group but given sampling recruitment problems this design had to be changed, as detailed below.

In order to allow a targeted and discipline-specific intervention a profile of the students' writing skills was developed. This profile directly fed into the goals that were to be elaborated in the intervention itself.

Sample

The intervention programme had been announced to all of the second year developmental psychology class at the start of the academic year (approximately 90 undergraduates) as part of a curriculum effort to improve essay writing in this psychology module. Our intervention group consisted of an opportunity sample of 23 undergraduates who registered for the intervention session. After trawling for more volunteers to attend intervention sessions 23 other undergraduates, who were re-contacted, acted as a non-intervention control. This group attended no sessions and was simply a comparative group.

All the undergraduates were second year undergraduate psychology students and were attending a course in developmental psychology as part of their psychology degree programme. They were aged between 19 and 35. The sample was primarily female with only 5 males out of the total of 46 participants. The two groups showed no significant difference in exam essay marks in a previous psychology exam.

Measures

Pre-intervention measure. The pre-intervention measure was a class test given to the whole second year undergraduate developmental psychology class. The aim of the class test was to provide students with some compulsory writing practice and tutor feedback in order to prepare them for their end of term written exam. The class test comprised a pre-

Table 1: Pre- and post-intervention measurement mean scores for the intervention and control group

Intervention group	Pre-intervention	SD	Post-intervention	SD
Essay tutor score	60.52	9.0	71.4	14.3
Essay word length	865 words	252.4	1154 words	360.7
Total rubric scores	61.2%	16.8	71.4%	14.3
Rubric introduction section	39.1%	23.9	42.9%	31.6
Rubric body overall section	67.6%	20.9	78.0%	15.8
Rubric within topics section	64.5%	16.7	73.1%	14.8
Rubric conclusion section	51.9%	29.6	69.4%	26.9
Control group	**Pre-intervention**		**Post-intervention**	
Essay tutor score	60.65	8.1	57.2	17.3
Essay word length	799 words	221.9	804 words	321.5
Total rubric scores	58.2%	16.4	57.9%	17.5
Rubric introduction section	33.7%	21.8	36.2%	27.2
Rubric body overall section	65.0%	17.3	65.9%	17.3
Rubric within topics section	61.2%	17.3	60.7%	19.0
Rubric conclusion section	48.2%	30.4	36.6%	31.0

seen essay task that was required to be completed in a one-hour timed session. The essay task was to write an essay to the prompt 'Defend the idea that humans are not born with an innate preference for faces'. Permission to access the essays was given when the students signed up for the study. Essays were available for the complete intervention and control groups. As shown in Table 1 no significant differences were found between the two groups in the results from this task.

The essays were marked by the course tutors, who included the first two authors, but all the markers were ignorant of who had signed up for the study. The marking criteria were the normal discipline specific criteria for the teaching course.

Post-intervention measure. The post-intervention measure was a pre-seen essay question answered in an exam with a recommended one-hour answer time. The exam was the second year developmental psychology exam and was taken eight weeks after the pre-intervention measure. The students had a choice of six questions from which to choose. The questions were all similar in format to the pre-intervention measure and demanded a persuasive essay response. Permission to access the essays was given when the students signed up for the study. The pre- and post-intervention essays were marked by the course tutors, who included the first two authors, but all the markers were ignorant of who had signed up for the study. The marking criteria were the normal discipline-specific criteria for the teaching course and as used for the pre-intervention essays.

Rating the essays

The essays were rated using two measures. First was the standard tutor mark for the piece of work as part of the developmental psychology course using discipline-specific criteria. The second rating was a quality rating using an essay evaluation rubric that we had designed based on criteria specified by Westby and Clauser (1999). This produced a measure of the profile-specific performance.

A rubric is a common method used to assess children's writing and can be adapted to assess all levels of writing. It uses a set of

rules or benchmarks to judge different levels of performance. The student's writing is assessed by its conformity to the benchmarks, in terms of a score or percentage measure of conformity. It is a measure of the structure and quality of the text produced by the students.

The rubric used in the current study consisted of 24 questions in four sections (see Appendix 1); introduction, overall body, within topics and conclusion. The rubric assessed students' skill at sectioning the essay clearly; ordering ideas; linking ideas; showing sufficient support and expansion of ideas and showing a sufficient sense of audience. The rubric displayed good internal reliability (Cronbach's alpha = .91). The item discrimination was high and it also showed good external validity (r = .79 with TOWL-3 standardised measure of writing). A copy of the rubric can be provided by the first author.

Procedure

All students in the second year developmental psychology class were offered the chance to participate in the intervention. This was done by announcements to the whole class and the use of posters in the psychology department. Those that signed up for the intervention were given a date and time to turn up for the intervention session, following their normal class towards the end of the semester.

The intervention was a one-hour small group lesson covering the achievement of structural subgoals in persuasive essay writing. These subgoals covered the following areas identified as weak in the students writing; sectioning the essay clearly (into introduction, topics, conclusion); ordering and linking ideas; showing sufficient support and expansion of ideas within and between topics.

The content for the lesson was derived from a previous study of 54 undergraduate developmental psychology essays drawn from the same cohort of students who were to be asked to take part in the intervention. This study revealed that the students were weak in a number of areas. They were very poor at producing good introductions and conclusions. For example, 51 per cent of this cohort of students did not explain in their introduction what was to be covered in the essay and 35 per cent of them did not produce a conclusion that summarised any of the main points of the essay. They were also poor at making links between topics and they failed to use adequate vocabulary to link between topics – only 18 per cent of student completely linked topics and 25 per cent of students did not use any linking words between topics. However, they did use evidence to illustrate points and 85 per cent of the students did produce topics in a logical manner related to the question. The essays produced by all the students were also very short, given the level of detail the students were expected to attain in their essay. We therefore designed our lesson to cover each of the perceived areas of weakness of the cohort we had identified from the profile above.

The lesson contained seven major sections as listed in Table 2. As previously out-

Table 2: The breakdown of the one-hour intervention lesson

Title of section	Time in minutes
1. Introduction and explanation of the genre and how lesson will run (use of their work, etc.)	5
2. Presentation of overall macrostructure of the persuasive essay	10
3. Presentation of introduction structure subgoal	10
4. Presentation of conclusion structure subgoal	10
5. Presentation of topics sections in essay subgoal	10
6. Presentation of linkage subgoal	10
7. Brief discussion of what had been learnt	5

lined these sections had been identified as weak areas from the profile of students' work. Each section of the lesson followed a common structure. To begin with the present knowledge and writing habits of the students were elicited. This was done using the pre-intervention measure they had been asked to bring with them to the intervention. A model subgoal was presented and there then followed a group discussion about this goal and how it could be applied. The students then applied the subgoal to their own work.

The emphasis was placed on interaction and reflection of the subgoals. All students were encouraged to participate equally throughout. The session was led by the third author who has extensive experience in tutoring psychology students' essay writing skills. The other authors were psychology tutors on the module teaching staff and so provided further discipline-specific guidance and representation in the design of the intervention session. Each section of the hour-long session is explained in more detail below.

1. Introduction and explanation of the genre and how lesson will run. Participants were asked to consult their pre-intervention class test essay. They were told that they would be referring to these throughout.

Students were informed of the reason for the study. They were told that university essays had been studied for structure. They were informed that many of these essays lacked the features deemed necessary for successful essay writing, according to the literature. Students were told that the aim of the intervention study was to make them aware of these features in their own writing and to see if instruction in this area might improve their grade at a later stage.

The students were informed that the essay writing genre was a specific and restricted genre governed by specific writing conventions. They were told that the intervention would be focusing on the structure of the persuasive essay genre in psychology. Participants were informed that they would first look at the overall structure of a psychology essay and then examine the individual parts in more detail and that finally they would be looking at cohesion – joining it all together.

For this and all subsequent sections the section ended with a summary by a group member, final discussion and a question and answer session.

2. Presentation of overall structure of the persuasive essay. Participants were asked if they felt that they structured their essays. This elicited a brief group discussion. Participants were told that an academic essay should have easily definable sections: an introduction, a main body and a conclusion. Participants were given an sample essay and asked to mark on it where they believed the section boundaries to be. This task was difficult, in cases, impossible, as the essay contained very little structure. This task then led on to another brief group discussion in which the importance of structure was drawn from the example used.

Students were asked to carefully consider their own work and identify sections in their own essay from the class test. This was followed by a brief group discussion.

3. Presentation of introduction structure. The tutor led a discussion on the purpose of the introduction and what an introduction should contain. Participants were alerted to the varying nature of the introduction section. Students were presented with a number of subgoals relating to the introduction structure and asked to identify and discuss features of the introduction in their own class test work. This was done in pairs and then as a group.

4. Presentation of conclusion structure. The tutor led a brief discussion on the purpose of the conclusion, what a conclusion should contain and what a conclusion should not contain (i.e. new information). Students were presented with a number of subgoals relat-

ing to the conclusion structure.

Participants were given two conclusions to study. They were asked to comment on them and decide which was the better structured of the two. The students then attempted to identify and discuss features of the conclusions in their own work. This was done in pairs and then as a group.

5. Presentation of topic sections in essay. The tutor led a discussion on how the body of an essay should be organised. The tutor argued that the main body of the essay should be broken down into 'topics'. Students were presented with a number of subgoals relating to topics. There followed a brief discussion of the structural features of each topic/section. The students then attempted to identify and discuss features of the topics in their own work. This was done in pairs and then as a group.

6. Presentation of linkage. The importance of ordering and linking information appropriately was stressed in a discussion led by the tutor. Linking words, connectives were defined and the group named as many as they knew. Students were presented with a number of subgoals relating to linkage. The students worked on an example essay identifying the connectives and linking words. Using their own work the students counted linking words in the two sections. This was done in pairs and then as a group.

7. Brief discussion of what had been learnt. Students were reminded of how the intervention was designed and how they might use the subgoals. The subgoals were then given to them on a handout. Students read through the handout, using the tutor to help define any unknown terms or words (e.g. vacillate).

Results

The pre-intervention and post-intervention essay measures and the rubric assessment marks are given in Table 1. The marks given for the rubric are expressed as a percentage of the summed maximum marks that the students could have received by section. There were no significant differences between the intervention group and the control group for any of these measures at pre-intervention. Therefore, it is reasonable to assume that the two groups were matched in their essay writing ability prior to the intervention.

It can be seen that the results of the intervention group were higher, in general, than the control group post-intervention. The scores of the two groups were compared with repeated measures analysis of variance.

Tutor marks and essay length

Analysis of variance on the pre- and post-intervention essay marks showed no significant main effects but did show a significant interaction between tutor mark and group (tutor mark, $F(1,44) = 3.6$, $p > .05$, $\eta = .75$; group, $F(1,44) = 2.8$, $p > .05$, $\eta = .6$; tutor mark with group, $F(1,44) = 4.6$, $p < .05$, $\eta = .88$). The post-intervention scores of the intervention group were significantly higher than those of the control group (planned comparison post-intervention t test, $t(44) = 2.69$, $p = .01$, $d = 0.8$).

As noted from the student writing profile, the students' essays in the pre-intervention were short for the level of detail expected within the essay itself. An ANOVA comparing pre- and post-intervention essay length count showed that the main effects were significant (essay length, $F(1,44) = 10.5$, $p < .01$, $\eta = .20$, group, $F(1,44) = 7.8$, $p < .05$, $\eta = .15$) and the interaction of essay length and group was significant ($F(1,44) = 6.7$, $p < 0.01$, $\eta = .18$). A planned comparison t test confirmed ($t(44) = 3.5$, $p < .01$, $d = 1.02$) that the intervention group produce longer post-intervention essays. There was a significant correlation between answer length and the final grade received across both groups ($r(46) = .44$, $p < .01$).

Analysis of essays by sections and by length

The parts of the essay were broken down by the raters into introduction, body and conclusion and a word count made of the length

of each of these sections. We were interested in comparing word counts by section as the initial profile of students work had shown a number of these sections to be extremely short or in many cases not there at all. Therefore, length provided a ready indication of how effective the intervention had been at getting the students to include some more detail in these sections. A repeated measures analysis of variance looking at the pre- and post-intervention word counts showed a number of statistically significant differences. When examining body length there were significant main effects (body, $F(1,42) = 11.7$, $p < .01$, $\eta = .2$; group, $F(1,42) = 4.8$, $p < .05$, $\eta = .1$) and there was a significant interaction between body length and group ($F(1,42) = 4.7$, $p < .05$, $\eta = .4$). Examination of the length of the conclusion section revealed a significant main effect of group and a significant interaction of conclusion length by group (conclusion length, $F(1,42) = 0.7$, $p > .05$, $\eta = .01$, group, $F(1,42) = 7.8$, $p < .05$, $\eta = .16$; conclusion length by group $F(1,42) = 9.1$, $p < .01$, $\eta = .18$). This showed that the intervention group produced significantly higher amounts of words in both the body and conclusion section of the post-intervention measure (body length planned comparison post-intervention *t* test, $t(44) = 2.97$, $p < .01$, $d = 0.88$, and conclusion length planned comparison post- intervention *t* test, $t(44) = 5.4$, $p < .01$, $d = 1.6$). However, there were no main effects ($F < 1$) or significant interaction ($F < 1$) between introduction word length and group or any other main effects ($F < 1$). The intervention group did not produce a longer introduction than the control group in the post-intervention measure and so gains were specific to the body and conclusions sections.

Analysis of essays by the rubric scores

The scores for the writing rubric were also assessed to see if there were any post-intervention improvements. Pre- and post-intervention scores are shown in Table 3 for the intervention group on all the questions. Analysis of variance on the total rubric scores showed that there was a significant main effect of group but not of total rubric scores (total rubric score, $F(1,44) = 2.5$, $p > .05$; $\eta = .06$, group, $F(1,44) = 4.9$, $p < .05$, $\eta = .10$) and that there was also no significant interaction between pre- and post- rubrics total scores and group ($F(1,44) = 3.8$, $p = .056$, $\eta = .08$). However, when we converted the scores into gain scores and covaried the baseline scores we did find a significant effect of group ($F(1,45) = 7.7$, $p < .01$, $\eta = .15$). There was also a significant relationship between essay length and score on the rubric that was retained, across both groups ($r(46) = .64$, $p < .01$).

It was also shown that once the rubric was broken down into the component sections, pre and post differences between the groups were shown for the conclusion section. Analysis of variance revealed that there was main effect of group (conclusion rubric score, $F(1,42) = 0.2$, $p > .05$, $\eta = .01$; group, $F(1,42) = 8.5$, $p < .05$, $\eta = .2$) and a significant interaction between pre- and post-intervention conclusion section rubric scores and group ($F(1,42) = 5.1$, $p < .05$, $\eta = .1$, conclusion section planned comparison post-intervention *t* test, $t(44) = 3.9$, $p < 0.01$, $d = 1.15$). The intervention group was scoring more highly on the conclusion section of the rubric than the control group in the post intervention measure. This result makes sense in terms of the very short average length of the conclusion section that the control group produced, in comparison with the significantly longer conclusion section produced by the intervention group.

The responses to individual questions were also analysed to see if there were pre and post differences between groups. The questions shown in Table 4 displayed significant interactions where there were differences between the intervention and control groups for post-testing and not for pre-testing. The results show that the students in the intervention group are using more linkage in their post-intervention work as recommended in the intervention. The intervention group are also more likely to include

Table 3: Intervention group pre and post intervention mean rubric scores by question

	Pre-intervention Mean rating (0–3)	SD	Post-intervention Mean rating (0–3)	SD
Introduction section				
Does information presented in the introduction show that the writer has identified terms in the title?	2.32	0.89	2.26	1.1
Does the introduction explain what is to be covered in the essay?	0.91	1.02	1.09	1.35
Does the introduction explain what order topics will be covered in the essay?	0.27	0.77	0.52	0.9
Body section				
Does the writer vacillate between topics?	2.55	0.74	2.48	1.08
Are topics logically ordered?	2.32	0.84	2.83	0.39
Are the topics linked to the question?	2.27	0.7	2.7	0.47
Are the topics linked to each other?	1.91	0.87	2.17	0.83
Has the writer made use of basic linking words between topics?	2.0	0.82	2.22	0.85
Has the writer made use of complex linking and referential words/ phrases/ sentences between topics?	1.05	0.84	1.3	1.18
Within topics section				
Are ideas relating to one topic logically ordered?	2.48	0.6	2.74	0.45
Does the writer include identifiable topic sentences?	2.09	0.81	2.43	0.79
Is the main idea expressed in the topic sentence developed?	1.73	0.7	2.04	0.47
Does the writer explain/define necessary/relevant terminology?	1.59	0.5	1.83	0.58
Has the writer made use of explanation/description to back up the main idea of the sections?	2.0	0.69	2.35	0.65
Has the writer made use of example/case studies/ evidence from research/data to back up the main idea of the topic?	2.18	0.73	2.3	0.63
Is the relevance of these illustrations explained to the reader?	1.68	0.78	2.13	0.63
Does the writer go in evaluating the illustrations presented?	1.73	0.83	2.0	0.6
Are the ideas about the topic linked?	2.45	0.69	2.52	0.59
Has the writer made use of basic linking words to link ideas within discussion of a topic?	2.27	0.63	2.43	0.59
Has the writer made use of complex linking words/ phrases/ sentences to link ideas relating to a topic?	1.09	0.75	1.35	1.03
Conclusion section				
Does the ending summarise the main points in the essay?	1.05	0.84	1.78	0.9
Are the summarising points and the concluding comments linked?	1.32	1.04	2.0	0.95
Does it refer back to the title?	2.05	1.0	2.48	0.85
Does it answer the question or explain why the question cannot be answered?	1.77	1.07	2.09	1.04

Table 4: Rubric question where a significant post-intervention difference was found between the intervention group and the control group

Within topics section
'Are the topics linked to the question?' (Interaction of question by group, $F(1,42) = 5.3$, $p = .027$, planned comparisons t test on post-intervention scores, $t(44) = 2.7$, $p = .01$, $d = 0.8$)
'Is the relevance of these illustrations (examples/case studies/evidence/data) explained to the reader?' (Interaction of question by group $F(1,42) = 5.04$, $p = .03$, planned comparisons t test on post-intervention scores, $t(44) = 2.9$, $p = .005$, $d = 0.9$)
Conclusion section
'Does the ending summarise the main points in the essay?' (interaction of question by group, $F(1,42) = 5.8$, $p = .021$, planned comparisons t test on post-intervention scores, $t(44) = 3.5$, $p = .001$, $d = 0.5$)
'Are the summarising points and the concluding comments linked?' (interaction of question by group, $F(1,42) = 5.3$, $p = .026$, planned comparisons t test on post-intervention scores, $t(44) = 4.2$, $p < .000$, $d = 1.2$)
'Does it (the conclusion) refer back to the title?' (interaction of question by group $F(1,42) = 4.1$, $p = 0.049$, planned comparisons t test on post-intervention scores, $t(44) = 3.8$, $p < .000$, $d = 1.1$)
'Does it (the conclusion) answer the question or explain why the question cannot be answered?' (interaction of question by group, $F(1,42) = 6.5$, $p = .015$, planned comparisons t test on post-intervention scores, $t(44) = 3.0$, $p = .005$)

supporting evidence in their essays. The overall improvement in intervention group conclusion scores described earlier is reflected in an overall improvement in all the specific conclusion questions in the rubric and they showed superior structuring and content of their conclusions.

Discussion

The discipline-specific profile of student weakness in writing led to the design of our discipline-specific intervention. This profiling before intervention was, we think, an important factor in the success of our following intervention. It allowed us to target the specific areas of student weakness in their psychology essays in a short space of time rather than spending more time on perceived generic problems with writing. This specificity allowed psychology tutors to improve the writing skills of their students within a psychology module and within a realistic time frame, given other curricular demands. The lesson could also be successfully imported into a larger package more oriented towards the WID approach.

Psychology tutors can influence the writing of their students through direct, explicit teaching in a way similar to that of Graham (in press) and colleagues with children in the USA. The writing lesson requires no great knowledge of the English language or specific remedial teaching skills and could be implemented, in our opinion, by most psychology tutors without an excessive resource cost to departments. Pragmatically, the lesson also sits well with psychology tutors who may be more ideologically opposed to the 'academic literacies' approach (Lea, 2004: Lea & Street, 2000) that is currently gaining ground in relation to the teaching of student writing.

The intervention targeted performance in one module in psychology where students

were profiled as poor at the structural organisation of their essays. The intervention provided students with information on the purpose of each section of the essay and identified subgoals relevant to the production of academic essays. The short intervention led to a number of improvements and had an effect on the final tutor mark awarded to the essays and also led to longer essays overall. The intervention group students linked their supporting material in the essay more to the question and they also explained more about why their supporting material was relevant to the question. The intervention students were also more likely to produce a clearly structured conclusion. Both the conclusion length and the conclusion rubric score produced large effect sizes. For the conclusion length the mean of the intervention group was at the 73rd centile of the comparison group and for the rubric conclusion score the mean of the intervention group was at the 62nd centile of the comparison group. These positive results support previous work in this area that places goal elaboration as central to a self-regulated model of successful writing (Graham, in press).

However, the intervention did not improve the introduction section of the essays analysed or students' ability to use complex linking words. This differential pattern of results indicates that there was not a generalised improvement in student performance but targeted effects on writing. This would argue against the idea that any effects found are due to a general placebo effect from the intervention. The lack of improvement could be due to problems in the intervention in these areas or could suggest that there are certain elements of the writing process that are more resistant to change than others and may require additional support and practice.

It is interesting to take this further and speculate why the introduction section did not improve after the intervention while the conclusion did. The introductions of all the writers in this sample were very brief and appeared to be rushed, despite the intervention group's tutoring in how to structure an introduction. Students in both groups were very keen to begin writing in the exam essay and so they may have postponed most of their planning until during the writing process itself. Assuming the intervention group planning process was influenced, somewhat, by the intervention instruction then they would apply the lessons learnt in the intervention session as they wrote. Hence, they would produce more clearly structured conclusions and use their knowledge to link topics and support evidence appropriately. As such the introduction would be less influenced by the intervention. In fact, Torrance (1996) reported that less than 15 per cent of essay writing time was taken up with planning activities in undergraduate essay writing tasks. This is in contrast to the established view that 'expert' adult writers devote considerable amounts of time to planning (Kellog, 1988). The study did not provide the opportunity to question the students about their post-intervention approaches to writing.

The intervention did not appear to lead to any greater frequency in the use of complex linking words. The use of basic and complex connectives has been argued by Freedman and Pringle (1994) to be important to ensure the synthesis and the chaining of the logical sequence of statements that are needed for persuasive essay texts to work. We saw no improvement in the use of basic or complex connectives in the intervention group despite significant correlations of frequency of use of basic and complex connectives with overall grades and rubric scores. However, there was evidence that basic connectives were being used and it was the complex connectives that were lacking. Therefore, despite being introduced to complex and less frequent connectives in the intervention their more frequent use may require more extended practice and feedback than we were able to give in our short lesson.

The main area where the intervention showed a clear success was in the improvement to the structure of the conclusion section of the essay. The intervention group improved significantly in their scores for all the rubric ratings relevant to the conclusion section. In psychology persuasive essays conclusions are very important. They provide the opportunity for writers to summarise their line of argument and make their final influence on the opinion of the reader (Tompkins, 1994). In that respect the intervention can be seen as a success, despite being a very short lesson to the students.

The intervention group essays were also considerably longer than those of the control group. The length of the pre-intervention essays had been rated by tutors as too short to contain the level of detail they expected on the topic itself. Our results would seem to support the idea that the intervention students had more cognitive resources available to devote to text production since they needed to devote less to essay structuring. Clear goals on how to structure the essay provided clear direction for the students in their compositions. Well-specified goals also provide motivation and improve self-regulation towards completing tasks since they make it possible to monitor progress towards the overall goal of answering the essay question (Ferretti *et al.*, 2000). More resources devoted to composition and a more general motivating and monitoring mechanism resulting from learning appropriate structural goals contributed towards the production of more words in the intervention group's essays.

Structural quality is not everything, though. Content and topic knowledge are the primary contributors to overall grades (Westby & Clauser, 1999) and this is probably why essay length correlates with essay grade. It is interesting to note that for the conclusion section of the post-intervention essay, five of the intervention group produced good conclusions with high overall structural scores on the rubric. However, their overall grade was not very high and their texts were shorter than the average. This probably reflects the contribution of content to their essay scores. They were able to produce an appropriately structured essay; but if it lacks content it will not receive high grades.

This was a small-scale study using a short intervention. However, despite this we achieved an increase in exam marks relative to a control group. This is all the more interesting since the lesson was only one hour in length. We carefully designed the intervention to match the profile of student weakness we had discerned from our rubric measure and so this allowed the intervention to be short but effectively targeted. Therefore, the effects of the intervention were specific, as a result of this profiling, and not general, so arguing against a placebo effect. However, clearly there is a need to further develop and further evaluate such intervention methods. We do not know if our evaluation had any longer-term effects nor if it was specific to exam essays rather than more general coursework essays. It could also be argued that since the intervention group students signed up to the study then they were probably the most motivated students and by showing their willingness to attend the intervention session were probably also the most open to new learning experiences. However, since the intervention was very specific and designed to be in line with other interventions that have had long-term effects on children (Graham, in press) then we are confident that the students will continue to make improvements in their writing.

What was very clear from the study was that our students needed support in their writing. Our profile of writing showed considerable student weakness. Researchers in writing are beginning to question the assumption that adults are necessarily 'skilled writers' (Johnstone *et al.*, 2002; Torrance, 1996). We may need to move away from the idea in higher education that writing tuition is somewhat remedial in nature and not for the 'normal' student (Hartley, 2001) towards a more involved and partici-

pative recognition that our students need to develop their writing skills throughout their time at university. This would ideally be within a larger-scale WID approach but, as we have demonstrated, can also be fitted into smaller-scale department-level initiatives.

Address for correspondence
Vincent Connelly, Department of Psychology, Oxford Brookes University, Gipsy Lane, Oxford, OX3 0BP.
E-mail: vconnelly@brookes.ac.uk

References

Ashbaugh, H., Johnstone, K.M. & Warfield, T.D. (2002). Outcome assessment of a writing-skill improvement initiative: Results and methodological implications. *Issues in Accounting Education, 17,* 123–148.

Bangert-Drowns, R.L., Hurley, M.M. & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74,* 29–58.

Bereiter, C. & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Lawrence Erlbaum.

Channok, K. (2000). Comments on essays: Do students understand what tutors write? *Teaching in Higher Education, 5,* 95–105.

Connelly, V., Campbell, S., MacLean, M. & Barnes, J. (2006). Contribution of lower-order letter and word fluency skills to written composition of college students with and without dyslexia. *Developmental Neuropsychology, 29,* 175–196.

Connelly, V., Dockrell, J. & Barnett, J. (2005). The slow handwriting of undergraduate students constrains overall performance in exam essays. *Educational Psychology, 25,* 97–105.

De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology, 97,* 139–156.

De La Paz, S. & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94,* 687–698.

De La Paz, S., Swanson, P.N. & Graham, S. (1998). The contribution of executive control to the revising by students with writing and learning difficulties. *Journal of Educational Psychology, 90,* 448–460.

Ferretti, R.P., MacArthur, C.A. & Dowdy, N.S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology, 92,* 694–702.

Freedman, A. & Pringle, I. (1984). Why students can't write arguments. *English Education, 18,* 73–84.

Graham, S. (in press). Writing. To appear in P.A. Alexander & P.H. Winne (Eds.), *Handbook of educational psychology.* Mahwah, NJ: Lawrence Erlbaum.

Hartley, J. (2001). Students, writing and computers. *Psychology, Learning and Teaching, 1,* 10–15.

Hayes, J.R. (1996). A new framework for understanding cognition and affect in writing. In C.M. Levy & S. Ransdell (Eds.), *The science of writing* (pp.1–27). Mahwah, NJ: Lawrence Erlbaum.

Hayes, J.R. & Flower, L.S. (1980). Identifying the organisation of writing processes. In L. Gregg & E.R. Sternberg (Eds.), *Cognitive processes in writing* (pp.3–30). Hillsdale, NJ: Lawrence Erlbaum.

Hayes, J.R. & Flower, L. S. (1986). Writing research and the writer. *American Psychologist, 41,* 1106–1113.

Hennessy, D. & Evans, R. (2005). Reforming writing among students in community colleges. *Community College Journal of Research and Practice, 29,* 261–275.

Hull, G. (1981). Effects of self-management strategies on journal writing by college freshman. *Research in the Teaching of English, 15,* 135–148.

Johnstone, K.M., Ashbaugh, H. & Warfield, T.D. (2002). Effects of repeated practice and contextual writing experiences on college students writing skills. *Journal of Educational Psychology, 94,* 305–315.

Kellog, R. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14,* 355–365.

Lavalle, E., Smith, J. & O'Ryan, L. (2002). The writing approaches of secondary students. *British Journal of Educational Psychology, 72,* 399–418.

Lea, M.R. (2004). Academic literacies: A pedagogy for course design. *Studies in Higher Education, 29,* 739–756.

Lea, M.R. & Street, B. (2000). Student writing and staff feedback: An academic literacies approach. In M.R. Lea & Stierer, B. (Eds.), *Student writing in higher education: new contexts* (pp. 32–46). Buckingham: Open University Press.

Maclellan, E. (2001). Assessment for learning: The differing perceptions of tutors and students. *Assessment and Evaluation in Higher Education, 26,*

307–318.

Matsuhashi, A. & Gordon, E. (1985). Revision, addition and the power of unseen text. In S. Freedman (Ed.), *The acquisition of written language: Response and revision* (pp.236–249). Norwood, NJ: Ablex.

McCune, V. (2004). Development of first-year students' conception of essay writing. *Higher Education, 47,* 257–282.

Monroe, J. (Ed.) (2002). *Writing and revising the disciplines.* Ithaca, NY: Cornell University Press.

Norton, L.S. & Norton, J.C.W. (2001). The essay feedback checklist: How can it help students improve their academic writing? Paper and workshop given at the first international conference of the European Association for the Teaching of Academic Writing across Europe (EATAW), Groningen, The Netherlands, 18–20 June 2001. ERIC abstract. No. ED45430.

Norton, L., Brunas-Wagstaff, J. & Lockley, S. (1999). Learning outcomes in the traditional coursework essay: Do students and tutors agree? In C. Rust (Ed.), *Improving student learning: Improving student learning outcomes.* Oxford: The Oxford Centre for Staff and Learning Development.

Nussbaum, M.E. & Kardash, C.A.M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology, 97,* 157–169.

Perkins, D.N. (1985). Reasoning as imagination. *Interchange, 16,* 14–26.

Singleton, C.H. (Chair) (1999) *Dyslexia in higher education: Policy, provision and practice.* Report of the National Working Party on Dyslexia in Higher Education. Hull: University of Hull on behalf of the Higher Education Funding Councils of England and Scotland.

Sommerville, E.M. & Creme, P. (2005). 'Asking Pompeii questions': A co-operative approach to Writing in the Disciplines. *Teaching in Higher Education, 10,* 17–28.

Riordan, D.A., Riordan, M.P. & Sullivan, C.M. (2000). Writing across the accounting curriculum: An experiment. *Business Communication Quarterly, 63,* 49–59.

Tompkins, G.E. (1994). *Teaching writing: Balancing process and product* (2nd edn). New York: Merrill.

Torrance, M. (1996). Case studies for familiar writing tasks: Case studies of undergraduates writing essays. In G. Rejlaarsdam, H. van den Bergh & M. Couzjin (Eds.), *Theories, models and methodology in writing research* (pp.283–298). Amsterdam: Amsterdam University Press.

Westby, C.E. & Clauser, P.S. (1999). The right stuff for writing: assessing and facilitating written language. In H. Catts & A. Kamhi (Eds.), *Language and reading disabilities* (pp.259–324). New York: Allyn & Bacon.