UNIVERSITY of
BRADFORD

Library

The University of Bradford Institutional Repository

http://bradscholars.brad.ac.uk

This work is made available online in accordance with publisher policies. Please refer to the repository record for this item and our Policy Document available from the repository home page for further information.

To see the final version of this work please visit the publisher's website. Access to the published online version may require a subscription.

Link to publisher version: *http://dx.doi.org/10.1155/2014/367493*

Citation: Mann J, Protty MB, Duffy J et al(2014) Near-Peer Teaching and Exam Results: The Acceptability, Impact, and Assessment Outcomes of a Novel Biological Sciences Revision Programme Taught by Senior Medical Students. Journal of Biomedical Education. 2014: 1-9.

Research Article

Near-Peer Teaching and Exam Results: The Acceptability, Impact, and Assessment Outcomes of a Novel Biological Sciences Revision Programme Taught by Senior Medical Students

Jake Mann,[1] Majd B. Protty,[1] John Duffy,[1] Mohammed Mohammed,[2] and Connie Wiskin[1]

[1] *College of Medical and Dental Sciences, University of Birmingham, Birmingham B15 2TT, UK*
[2] *Faculty of Health Studies, University of Bradford, Bradford BD7 1DP, UK*

Correspondence should be addressed to Majd B. Protty; majd@protty.net

Background. Near-peer teaching is becoming increasingly popular as a learning methodology. We report the development of a novel near-peer biological sciences revision course and its acceptability and impact on student confidence and exam performance. *Methods*. A cross-sectional analysis of tutee-completed evaluation forms before and after each session was performed, providing demographic details, quality scores, and self-rating of confidence in the topic taught on a 0 to 100 mm visual analogue scale (VAS). The confidence data was examined using analysis of means. Exam performance was examined by analysis of variance and canonical correlation analysis. *Results*. Thirty-eight sessions were delivered to an average of 69.9 (±27.1) years 1 and 2 medical students per session generating 2656 adequately completed forms. There was a mean VAS gain of 19.1 (5.3 to 27.3) in self-reported confidence. Looking at relationship between attendance and exam scores, only two topics showed significant association between number of sessions attended and exam performance, fewer than hypothesised. *Conclusion*. The present study demonstrates that near-peer teaching for biological sciences is feasible and is associated with improved self-reported confidence in the sessions taught. The outcome data, showing significant effect for only a small number of items, demonstrates the difficulty of outcome related research.

1. Introduction

Peer assisted learning is a well-established methodology, practised across many educational disciplines. Several derivations of both terminology and teaching methods have developed [1]. Near-peer teaching—used here—is defined as teaching delivered by a trainee who is one or more years senior to another trainee on the same medical training course [2, 3], in this report, senior medical students teaching junior medical students.

A systematic review of peer teaching and learning in clinical education in 2007 concluded that "peer teaching and learning is an effective educational intervention for health science students on clinical placements" [4]. In medicine near-peer teaching has been employed across a range of topics varying from anatomy to diversity awareness [5] to patient interviewing [6]. In line with the General Medical Council imperative relating to the doctor "developing the skills and practices of a competent teacher" teaching is considered good practice as part of students' professional development [7]. A review of the literature [8] concluded, interestingly, that despite peer teaching offering well-defined advantages and being widely employed, little is published. This may be because of difficulties associated with educational intervention meeting the outcome based preferences of peer review. The authors above concluded that "the analogy of the "journeyman," as intermediate between "apprentice" and "master," with both learning and teaching tasks, is a valuable but yet underrecognised source of education in the medical education continuum." This has echoes of Dreyfus and Dreyfus' model of skill acquisition [9], as the near-peer teacher is moving from a stage of abstract learning to developing their own experiences of teaching.

In this study we utilise self-assessment in conjugation with summative examination marks to evaluate a peer-lead revision programme. Self-assessment is not a novel

concept [10] and is widely used at both an undergraduate (preclinical) level and undergraduate (clinical) level. As a research tool in medical education, the reliability of self-assessment is dependent upon the following: the quality of the study [10], the ability of the study population [11], and their familiarity with self-assessment [12]. It is also particularly useful for gauging abstract concepts such as "satisfaction" or "confidence" [13], which are difficult to evaluate by other means.

This paper seeks to add to the literature by including both evaluative data collected from a large cohort of recipients and outcome measures correlating attendance at extracurricular near-peer events with exam performance. As Whitman and Fife [3] remind us "to teach is to learn twice," so the benefits to learner and student teacher are worthy of highlight.

2. Methodology

2.1. Course Design and Evaluation. Three senior students in the top decile of their medical studies at the University of Birmingham (UoB) volunteered to design and deliver a total of 38 biological sciences revision sessions (76 hours) to Years 1 and 2 students ahead of their end of year examinations. The course was offered free of charge on an optional basis to all students ($n = 396$ in Year 1 and $n = 362$ in Year 2) during their spring (May 2011) revision period, two weeks before their scheduled final examinations. The revision sessions covered anatomy, physiology, histology, cell biology, biochemistry, immunology, general pathology, and pharmacology, closely-mirroring the formal curriculum of the UoB Medical School. Thirty-one of these sessions were delivered in a didactic 2-3 hour lecture-based format with predetermined learning outcomes. Seven "exam technique (ET) sessions" (one hour each) were delivered in a smaller group environment in a participatory format that did not have specific learning outcomes but which focused on discussion of commonly tested topics.

We elicited feedback from participating students before and after each session on anonymised self-completed evaluation forms. These asked attendees to rate the perceived impact of the teaching, its acceptability, perceived changes in confidence with the subject field, and their view on who should deliver this revision, peers, near-peers, or staff. The aim was to collect feedback to inform the ongoing development of the revision curriculum. After the end of year examinations a quantitative data analysis was undertaken to investigate the relationship between attendance at one or more revision class and outcome for the student, that is, their awarded examination scores. Near-peer teachers (MP and JM) undertook analysis of the questionnaires, while a senior academic (CW)—for obvious reasons of examination data protection—undertook analysis of the outcome data. Both data set analyses were supported by senior staff in medical statistics.

2.2. Statistical Analysis

2.2.1. Course Evaluation and Confidence Score Analyses. The feedback evaluation forms were collected after each session

and were entered onto a spread sheet for analysis. Only forms that contained at least the before and after self-confidence ratings were considered valid for the purposes of this study. Statistical analyses were performed using IBM SPSS Statistics 19 (IBM, USA) and Minitab (Minitab Inc, USA).

Continuous variables (age, self-reported confidence scores, and Likert scale responses) were summarised using means and standard deviations (SDs), whilst categorical variables (sex, year of study, and ethnicity) were summarised using counts and percentages.

Statistical analysis of confidence scores by session was undertaken using analysis of means (ANOM) [14]. Analysis of means (ANOM) is a well-established technique in industrial quality improvement that compares the mean scores of sessions with the grand mean (where the null hypothesis is no difference) and produces a graphical display of the groups with the aberrant groups appearing outside statistical limits (set at the 5% level in our study that is, $P \leq 0.05$), thereby aiding the assessment of statistical and clinical significance whilst controlling for multiple testing. Two ANOM analyses were undertaken. (1) For each session, the mean self-reported confidence scores before the revision sessions were analysed and plotted using ANOM with 5% limits to assess which sessions had significantly different confidence levels. (2) For each session the gain in confidence (=confidence score after − confidence score before) was analysed and plotted using ANOM with 5% limits to assess which sessions have significantly different gain scores. Minitab (Minitab Inc, USA) was used for ANOM calculations.

2.2.2. Free Text Comment Analysis. Two authors (MP and JM) reviewed all the written free-text comments reported by students and categorised them as generally indicating overwhelmingly positive/uncritical, neutral/mixed, or over-whelmingly negative/critical comments via a thematic analysis. Five comments for each category were purposefully selected and presented for illustration.

2.2.3. Exam Score Outcome Analysis. For each topic and year students were classified according to the number of tutorials attended and their results compared using one-way analysis of variance. Near-peer tutorials were offered on a range of topics falling in the general area of biological sciences, but other topics were also examined (generally in the social sciences). Canonical correlation analysis was used to examine potential relationships between scores in these two areas.

2.3. Ethical Approval. Ethical approval was obtained internally at UoB by the University Ethics Review Committee which classified the project as educational evaluation (routine when validating any new teaching programme).

2.4. Hypotheses. The hypotheses were as follows.

(1) Revision class attendance would be acceptable to the student cohort and that they would perceive increased confidence in biological sciences topics as a result of the intervention.

TABLE 1: Profile of the responders to the student evaluation forms.

Age [mean (SD)]	19.4 (1.0)
Sex [male]	972 (37.6%)
Ethnicity [n (%)]	
Asian	555 (21.8%)
Black	30 (1.2%)
Chinese	74 (2.9%)
Mixed	111 (4.4%)
Other	41 (1.6%)
White	1733 (68.1%)
Confidence before session (VAS) [mean (SD]	42.2 (18.4)
Confidence after session (VAS) [mean (SD)]	61.3 (17.9)
Learning objectives were made clear (Likert) [mean (SD)]	4.6 (0.9)
There was enough time to ask questions (Likert) [mean (SD)]	4.6 (0.7)
Tutor appeared confident (Likert) [mean (SD)]	4.7 (0.5)
Overall quality of session (VAS) [mean (SD)]	83.8 (13.5)
Total number of evaluation forms	2656

(2) Relationship would be identified between attendance at revision and higher awarded examination scores for some topics. We speculated that the effect would be more marked in Year 1 (due to unfamiliarity with and more limited experience of/exposure to biological sciences topics)

(3) A relationship might be identified between the topics examined that near-peer revision classes were offered on (broadly "biological sciences") and other examined topics that the revision classes were not offered on (broadly "social science").

3. Results

3.1. Results 1: Course Evaluation and Confidence Score Analyses. We delivered 38 revision sessions to an average of 69.9 students per session (standard deviation 27.1, range from 30 to 126) that yielded 2656 adequately completed evaluation forms. Twenty-seven forms were discarded because they were incomplete. Table 1 shows the demographics of students collected on the valid evaluation forms. The overall mean age of students was 19.4 (SD 1.0, range from 19 to 32) of whom 37.6% were male and just over two-thirds were Year 1 medical students (67.5%), consistent with the year group make up. Overall, the majority of students reported their ethnicity as white (68.1%) and one-fifth as Asian (21.8%) with Chinese, mixed, and other race students making up the rest (10.1%).

The average overall quality score rating was 83.8 (range from 66.2 to 90.3) on a 0 to 100 mm visual analogue scale (VAS). Whilst most of the 38 sessions were consistent with this high rating (Figure 1(a)), two sessions (*Musculoskeletal*
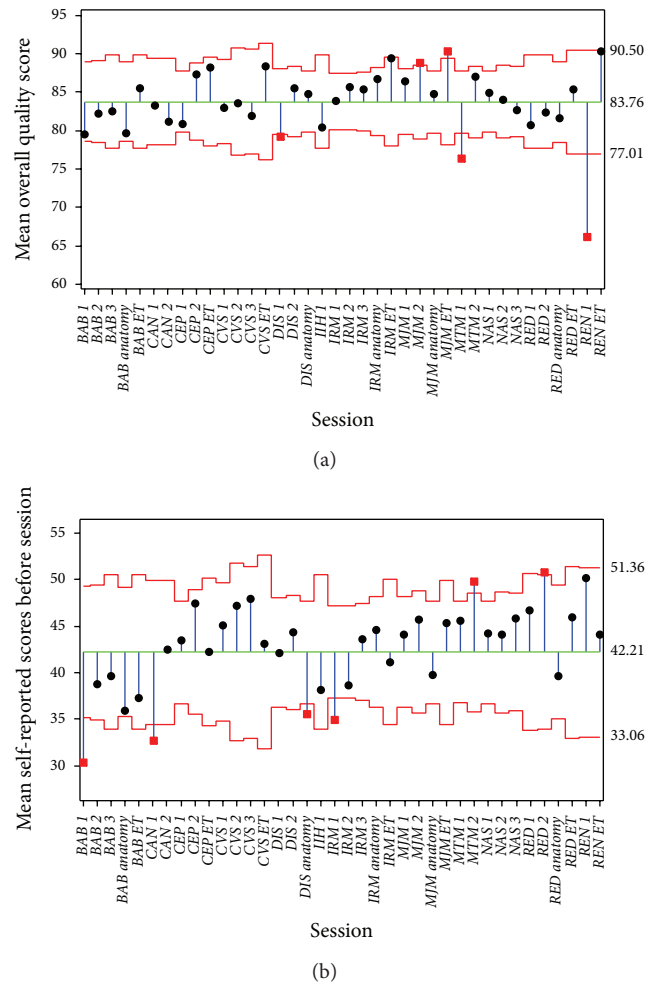

(a)

(b)

FIGURE 1: (a) ANOM of mean overall quality scores for the sessions. (b) ANOM plot showing mean self-reported confidence scores before the revision sessions. Stepped lines are upper and lower limits for 5% statistical significance.

(MJM) 2 and *MJM Exam Technique (ET))* were rated significantly higher and three sessions (*Cell biology (MTM) 1*, *Digestive (DIS) 1*, and *Renal (REN) 1*) were rated significantly lower than the average overall quality rating ($P < 0.05$). The evaluations from each session on a 5-point Likert scale were overwhelmingly positive, with students praising each session in a number of areas, including clarity of learning objectives, time dedicated for questions, and whether the tutor appeared competent, with an average rating of 4.6 (SD = 0.9), 4.6 (SD = 0.7), and 4.7 (SD = 0.5), for each characteristic, respectively.

The overall average self-reported confidence before the sessions was 42.2 (range from 30.3 to 50.8) on a 0 to 100 mm VAS. Whilst most of the 38 sessions were consistent with this rating (Figure 1(b)), two sessions (*Musculoskeletal 2* and *Reproductive and Development (RED) 2*) were rated significantly higher and four sessions (*Neuroscience (BAB) 1*, *Cancer (CAN) 1*, *Digestive anatomy*, and *Respiratory (IRM) 1*) were rated significantly lower than the average self-reported rating of confidence in the topic before the session ($P < 0.05$).

TABLE 2: Analysis of free text comments received on student evaluation forms.

	Type of comment		
	Overwhelmingly positive/noncritical	Neutral/mixed	Overwhelmingly negative/critical
Total number of comments	41/98 (42%)	41/98 (42%)	16/98 (16%)
Example 1 [*sic*]	"*Lots of time to ask questions well explained.*"	"*A lot of information that i didn't know, not suggesting that teaching was bad.*"	"*Too much to fit into given time. Maybe longer session or split into two.*"
Example 2 [*sic*]	"*Made me realise how to focus my revision and what to focus on*"	"*Allow fewer Questions to be asked.*"	"*Too long, hard to concentrate. Maybe have some sessions aimed at confident people and have some aimed at people who haven't got a clue*"
Example 3 [*sic*]	"*Really helpful + patient. Thanks*"	"*Have these sessions all year round.*"	"*A bit fast*"
Example 4 [*sic*]	"*Excellent, on previous forms i made clear that my confidence levels decreased. This is only because it was made clear how much i do not know.*"	"*It would be good if the bigger diagrams were printed on separate papers so that they were longer & cleaner.*"	"*Found it hard to follow because a lot of knowledge was assumed to already be known.*"
Example 5 [*sic*]	"*Brilliant!:) Good job*"	"*When asking audience questions repeat the response of the person answering, the whole lecture theatre cannot hear everything the lecturer can without a microphone*"	"*Bit too detailed?*"

The average gain in self-reported confidence, calculated as the difference between self-reported confidence levels before and after each session, was 19.1 (range from 5.3 to 27.3). Whilst most of the 38 sessions were consistent with this positive gain in self-reported confidence (Figure 2), three sessions (*Digestive anatomy*, *Respiratory 1* and *Musculoskeletal 2*) had a significantly higher and three sessions (*Neuroscience ET*, *Cell biology 1*, and *Renal 1*) had a significantly lower than the mean gain in self-reported confidence compared to the mean ($P < 0.05$).

The overwhelming majority (85.0%) of students ranked near-peer teachers as their preferred tutors to deliver these comprehensive revision sessions, with 13.3% preferring faculty staff members and 1.7% preferring peer teachers.

3.2. Results 2: Free Text Comment Analysis. Ninety-eight forms across all sessions included free-text comments. These comments were thematically analysed and classified as in the main positive/uncritical, neutral/mixed, or overwhelmingly negative/critical. For each category, five example comments are reported to illustrate the main themes fed back by students. There were 41 (42%) positive comments, 41 (42%) neutral comments, and 16 (16%) negative comments (Table 2).

3.3. Results 3: Exam Score Outcome Analysis. Three sets of analyses were undertaken:

(a) outcomes for Year 1 examination candidates,

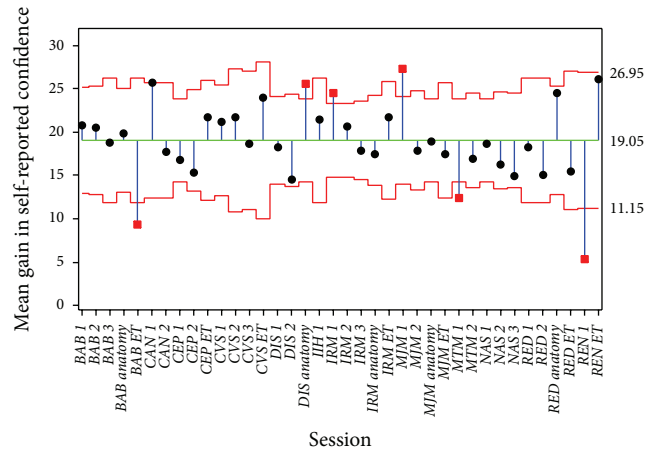(b) outcomes for Year 2 examination candidates,

FIGURE 2: ANOM plot showing mean gain self-reported confidence scores after the revision sessions. Stepped lines are upper and lower limits for 5% statistical significance.

(c) outcomes for relationships between the Modules examined in Year 1 and in Year 2.

3.3.1. Results 3(a): Outcomes for Year 1 Examination Candidates. Examination marks for each of the subjects for which tutorial were provided and were analysed by means of one-way ANOVA with number of relevant tutorials attended as the classifying factor (Table 3(a)).

Table 3: Analysis of variance for exam outcomes.

(a)

Number of sessions attended	Cell biology (MTM)				Cell signalling and Endocrinology (CEP)				Neurobiology (NAS)				Musculoskeletal (MJM)				Respiratory (IRM)				Digestive (DIS)			
	Mean mark	n	SD	P	Mean mark	n	SD	P	Mean mark	n	SD	P	Mean mark	n	SD	P	Mean mark	n	SD	P	Mean mark	n	SD	P
0	59.52	232	10.95	—	62.31	245	11.71	—	57.06	243	11.85	—	56.92	220	14.03	—	57.46	196	13.84	—	58.40	247	13.29	—
1	59.80	71	8.88	—	63.15	83	8.23	—	55.38	50	11.07	—	56.84	38	11.51	—	59.48	66	11.06	—	58.07	118	11.17	—
2	59.19	63	9.22	—	66.70	47	9.94	—	56.06	77	9.94	—	58.82	50	11.27	—	61.62	61	10.80	—	—	—	—	—
3	—	—	—	—	—	—	—	—	—	—	—	—	59.01	34	11.64	—	60.98	42	10.75	—	—	—	—	—
4	—	—	—	—	—	—	—	—	—	—	—	—	61.26	27	10.48	—	—	—	—	—	—	—	—	—
5	—	—	—	—	—	—	—	—	—	—	—	—	52.48	3	21.90	—	—	—	—	—	—	—	—	—
Total	59.52	366	10.27	0.89	63.05	375	10.89	**0.02**	56.62	370	11.36	0.40	57.64	372	13.06	0.12	58.93	365	12.63	**0.02**	58.30	365	12.63	0.82

(b)

Number of sessions attended	Renal (REN)				Cardiovascular (CVS)				Neuroscience (BAB)				Cancer (CAN)				Reproductive and Development (RED)			
	Mean mark	n	SD	P	Mean mark	n	SD	P	Mean mark	n	SD	P	Mean mark	n	SD	P	Mean mark	n	SD	P
0	61.78	275	10.56	—	58.67	292	10.07	—	60.58	237	11.50	—	59.93	271	9.82	—	67.63	287	8.73	—
1	62.38	51	8.35	—	57.83	44	10.47	—	59.84	36	12.43	—	61.22	49	8.88	—	67.08	47	8.88	—
2	61.15	31	11.71	—	58.22	21	7.91	—	65.50	28	10.11	—	62.08	38	11.01	—	67.47	24	8.82	—
3	—	—	—	—	—	—	—	—	62.47	36	11.02	—	—	—	—	—	—	—	—	—
4	—	—	—	—	—	—	—	—	59.96	21	8.90	—	—	—	—	—	—	—	—	—
5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Total	61.81	357	10.35	0.92	58.54	357	9.99	0.67	61.05	358	11.35	0.33	60.33	358	9.83	0.15	67.55	358	8.74	0.79

There is statistical evidence of a positive linear relationship between the number of tutorials attended and examination mark for *Cell signalling and Endocrinology (CEP)* (*P* value for linearity = 0.018) and *Respiratory (IRM)* (The *P* value for linearity = 0.016). There is no statistical evidence of relationship between the number of tutorials attended and examination mark for *Cell biology (MTM)*, *Neurobiology (NAS)*, *Musculoskeletal (MJM),* and *Digestive (DIS)*.

In summary, only two of the six topics offered generated data showing significant positive impact on examination score for those who attended revision classes.

3.3.2. Results 3(b): Outcomes for Year 2 Examination Candidates. Examination marks for each of the subjects for which tutorials were provided were analysed by means of one-way ANOVA with number of relevant tutorials attended as the classifying factor (Table 3(b)). The significance values provided in the tables relate to the linear component of the between-groups comparison; that is, the values provide a test of increasing mean score with increasing number of tutorials attended.

The numbers of tutorials attended was not associated with examination results for any of the topics.

3.3.3. Results 3(c): Outcomes for Relationships between the Modules Examined in Year 1 and in Year 2

Year 1

Canonical Correlation Analysis. The relationship between scores on the above topics (the end of year biological sciences examination topics which the senior students offered revision classes on) and the remaining topics (the *other* end of year examination topics which did not have revision classes offered) was explored by means of canonical correlation analysis.

Community Based Medicine 1 (CBM1), People, Patients and Populations (PPP), Doctors, Patients and Society (DPS), Integrated problems 1 (IP1), and Learning Medicine (LEM) are exams for the "social medicine" orientated part of the course.

The correlation matrix was constructed for the scores in these areas and canonical correlations calculated. The largest canonical correlation was 0.79, and the test for remaining correlations gave a chi-squared value of 22.5 on 20 df, indicating that the first pair of canonical variates accounts for the relationship between the two topic areas.

We are therefore only interested in the first pair of canonical variates. The standardised coefficients provide measures of the nature of the relationship between the two sets (Table 4). So in set 1 *Cell signalling and Endocrinology (CEP)* and *Neurobiology (NAS)* are not strongly related to the second set of variables, while in set 2 *Community Based Medicine 1 (CBM1)* and *Learning Medicine (LEM)* are not strongly related to the scores in the first set.

We conclude that the relationship between the two sets of scores is principally a relationship between the scores on *Cell biology (MTM), Musculoskeletal (MJM), Respiratory (IRM),* and *Digestive (DIS)* in the first set and *People, Patients and*

TABLE 4: Standardised canonical coefficients for the canonical correlation analysis.

Year 1	
Set 1 (biological sciences)	Canonical coefficient
Cell biology (MTM)	−0.305
Cell signalling and endocrinology (CEP)	0.077
Neurobiology (NAS)	−0.038
Musculoskeletal (MJM)	−0.213
Respiratory (IRM)	−0.302
Digestive (DIG)	−0.330
Year 1	
Set 2 (social sciences)	Canonical coefficient
Doctors, Patients and Society (DPS)	−0.680
People, Patients and Populations (PPP)	−0.272
Community Based Medicine 1 (CBM1)	−0.074
Integrated Problems (IP1)	−0.152
Learning Medicine (LEM)	−0.019
Year 2	
Set 1 (biological sciences)	Canonical coefficient
Immunity and Infection (IIH)	−0.145
Cardiovascular (CVS)	−0.319
Renal (REN)	−0.230
Neuroscience (BAB)	−0.065
Reproductive (RED)	−0.292
Cancer (CAN)	−0.098
Year 2	
Set 2 (social sciences)	Canonical coefficient
Health Services (HES)	−0.388
Decision Making (DEM)	−0.276
Community Based Medicine 2 (CBM2)	−0.095
Integrated Problems 2 (IP2)	−0.447
Student Project 1 (SP1)	−0.145

Populations (PPP), Doctors, Patients and Society (DPS), and (less strongly) *Integrated Problems 1 (IP1)* in the second set.

Year 2. A similar analysis for Year 2 students was performed. *Community Based Medicine 2 (CBM2), Health Services (HES), Decision Making (DEM), Integrated problems 2 (IP2),* and *Student Project 1 (SP1)* are exams for the "social medicine" orientated part of the course.

The analysis again resulted in a strong and highly significant first canonical correlation of 0.701 ($P < 0005$) with the test for remaining correlations giving a nonsignificant value of 22.7 on 20 df, indicating that the first pair of canonical variates accounts for the relationships between scores in the two topic areas. The standardised coefficients for the first pair of canonical variates are given in Table 4.

The relationship between the two sets therefore does not involve *Neuroscience (BAB)* or *Cancer (CAN)* very strongly with *Cardiovascular (CVS), Reproduction and Development*

(RED), and *Renal (REN)* doing most of the "work" in the correlation between the 2 groups of variables. For the 2nd set of variables most of the correlation with the first set comes from *Integrated Problems 2 (IP2), Health Services (HES),* and *Decision Making (DEM). Community Based Medicine 2 (CBM2)* contributes little to the relationship.

We conclude that the relationship between the two sets of scores is principally a relationship between the scores on *Cardiovascular (CVS), Reproduction and Development (RED),* and *Renal (REN)* in the first set and Integrated *Problems 2 (IP2), Health Services (HES), Decision Making (DEM),* and (less strongly) *Student Project 1 (SP1)* in the second set.

4. Discussion

The study elicited, encouragingly, that near-peer delivered revision sessions were highly rated by students and were associated with considerable gain in mean self-reported confidence scores. Looking at relationship between attendance and exam scores fewer significant associations were found than we hypothesised. These findings are discussed in turn below.

4.1. Evaluation. The perceived value of near-peer courses is two-fold. Firstly, these courses improve the teaching and leadership skills of the near-peer teachers. Given the demands of future doctors to serve as educators for both junior colleagues and patients, near-peer teaching during medical school is potentially an important curricular consideration [2]. Secondly, near-peer courses improve the confidence of tutees in their knowledge of the taught topics, as demonstrated by the results of our study. It is generally accepted that revision should boost confidence (although the reverse could be true where weakness is highlighted). While a limitation is not being able to compare confidence to nonattendees, it remains useful to know how well-received the initiative was.

The use of students to teach other students is implemented in many higher education institutes for medical and nonmedical education [1]. However, this model has not been formalised in teaching curricula nor routinely evaluated, which would be the case for teaching sessions delivered by university staff. The use and effectiveness of peer and near-peer teaching models in medical education have been described for communication skills [6, 15, 16], physical examination skills [17–19], and clinical skills training [20]. However, there is very little data available on near-peer teaching of undergraduate biological sciences revision sessions and their effectiveness and acceptance by tutees. As far as we know, this is the first study of this scale showing the feasibility and acceptability of near-peer teaching in delivering a comprehensive revision course for all biological sciences taught in the junior years of medical school.

Some sessions were significantly different from the others in self-confidence (Figures 1 and 2). These sessions merit further study. Sessions with significantly lower scores may reflect challenging content and/or poor delivery of sessions. Sessions with significantly higher confidence scores may reflect less challenging content or exemplar sessions which may provide insight for improving the delivery of other sessions. Sessions with significantly lower scores before the revision classes are likely to reflect the core degree teaching and highlight areas which may need improving or which students find difficult.

4.2. Relationship to Outcomes. Measurable impact was less than originally hoped for, but nonetheless the findings prompt reflection. There are difficulties with outcome based educational research, namely, (a) establishing which of a large number of existing and potentially influencing variables actually impacted on the examination score and (b) having the "blind spot" of not knowing how individual students *would* have performed had they not attended the intervention classes.

This may account at least in part for the dearth of reference in the literature to peer teaching, with authors reluctant to present reports that do not show evidenced "improvement." Much work is therefore unseen, making it particularly important that developmental work is disseminated.

It is important to acknowledge that multiple variables impact exam performance (home revision techniques/endeavour, individual tutor input, maturation, written sources, and exam nerves) so an examination score outcome can rarely be attributed to any single intervention or factor. It is equally important to acknowledge that lack of significant score improvement performance might mean *either* that the intervention had no impact, *or* that the intervention had an impact that the score data simply cannot show us. As an obvious example, while we "know" numerically that some students attended 6 hours of extra revision and, statistically, did no better than their peers who attended no extra revision classes, we do not know whether those students attending revision would have received lower scores than their peers had they not attended the classes. If, hypothetically, the revision classes (being self-selected and nonmandatory) attracted students who were weaker or less confident in those topics then their achieving the same results as their more confident and capable peers would, in fact, be a positive effect. To establish this effect we would need to either make the revision classes compulsory for all students to see if there was an overall improvement for everyone, correlate performance to previous marks in that topic to see if "struggling" students pulled up their marks, or do a randomized controlled trial (RCT). Mapping of previous performance would be possible for Year 2 but not Year 1 (this being their first attempt at a biological sciences test), and the RCT has ethical and logistic implications if some candidates are advantaged with support and others deprived. There is no obvious answer, but the discussion had to take this subject forward as an academic community.

In evaluating our near-peer revision programme we convincingly demonstrated an increase in self-assessed confidence. However, to draw on Kirkpatrick's hierarchy of learning evaluation [21], we did not universally demonstrate an improvement in examination performance. For reasons stated above, akin to other studies, it is difficult to conclusively prove a change in behaviour or results; this is particularly the case for such "theoretical" learning content. Performing

a RCT would potentially allow for higher level evaluation but would be challenging.

Additionally, score outcomes are arguably not absolute measures of teaching success. In the near-peer context here the very positive reported confidence increase should not be underestimated, and there are other potential advantages that resist traditional measurement. These include building relationships between senior and younger students, offering positive role modelling, professional development for 13 those delivering the classes, honed revision techniques for attendees, and so forth.

Though the participants gained in confidence, it was not always mirrored in their examination marks. This is not an uncommon finding [13] it is generally reported that higher performing students tend to underestimate their ability, whilst the converse is true for weaker students [11, 22]. It is difficult to estimate the magnitude of this effect because, as described above, we do not know how the students would have performed if they did not attend the revision courses.

Taking Results 3(a) and Results 3(b) at face value we can, and will, highlight classes that appear to have a significantly positive impact on examination performance and revise classes that do not. A clear area for review is the lack of relationships established in Year 2. The teachers were consistent between years, so we might usefully scrutinise the content and level and establish focus groups to better understand the type and needs of students attracted to these sessions.

Finally, considering Results 3(c), there were interesting findings from a curriculum point of view. In Year 1 CBM (*Community Based Medicine*) and LEM (*Learning Medicine*) did not have an outcome relationship with the other topics. While that might have been anticipated for LEM (being about learning methods), it was not anticipated for CBM which is based on general practice. Likewise there was not an association with IP (integrated problems). The important point is that both of these courses demand *integration* of aspects of science and social medicine. It is interesting therefore that they had weaker associations with results for "biomedical science" topics and "social medicine" topics than those two had with each other. In summary "science" and "social" results better predicted each other than the integrated topics did with either set. The relationship identified between the two separate fields, biomedical science and social science, is interesting given that students (from experience) tend to regard them as distinct. It may be of course that the brightest students excel across the board, but this initial finding does prompt discussion about the way modules relate to each other.

5. Conclusion

The present study demonstrates that near-peer teaching for biological sciences is feasible and is associated with improved self-reported confidence in the sessions taught. The outcome data, showing significant effect for only a small number of items, demonstrates the difficulty of outcome related research.

Further studies need to be undertaken to determine the value of such models in core medical curricula and their effectiveness in enhancing the performance of students in formal (summative) assessments.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
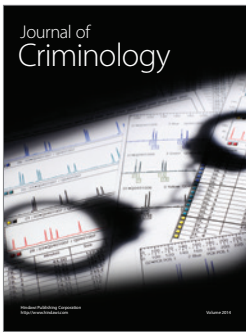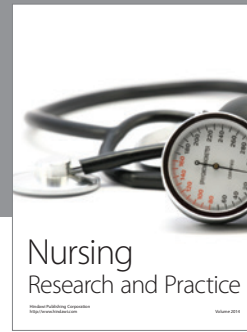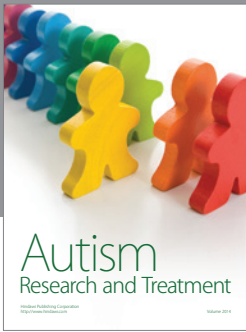
Acknowledgments

References

[1] D. C. Boud, R. Cohen, and J. Sampson, *Peer Learning in Higher Education*, Kogan, London, UK, 2001.

[2] C. Bulte, A. Betts, K. Garner, and S. Durning, "Student teaching: views of student near-peer teachers and learners," *Medical Teacher*, vol. 29, no. 6, pp. 583–590, 2007.

[3] N. A. Whitman and J. D. Fife, "Peer teaching: to teach is to learn twice," ASHE-ERIC Higher Education Report no. 4, Association for the Study of Higher Education, ERIC Clearinghouse on Higher Education, Washington, DC, USA, 1988.

[4] J. Secomb, "A systematic review of peer teaching and learning in clinical education," *Journal of Clinical Nursing*, vol. 17, no. 6, pp. 703–716, 2008.

[5] T. S. Tang, E. J. Hernandez, and B. S. Adams, ""Learning by teaching": a peer-teaching model for diversity training in medical school," *Teaching and Learning in Medicine*, vol. 16, no. 1, pp. 60–63, 2004.

[6] D. Nestel and J. Kidd, "Peer assisted learning in patient-centred interviewing: the impact on student tutors," *Medical Teacher*, vol. 27, no. 5, pp. 439–444, 2005.

[7] The General Medical Council, *Doctors as Teachers*, 1999.

[8] O. ten Cate and S. Durning, "Peer teaching in medical education: twelve reasons to move from theory to practice," *Medical Teacher*, vol. 29, no. 6, pp. 591–599, 2007.

[9] S. E. Dreyfus and H. L. Dreyfus, *A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition*, Storming Media, Washington, DC, USA, 1980.

[10] N. Kalchikov and D. Boud, "Student self-assessment in higher education: a meta-analysis," *Review of Educational Research*, vol. 59, pp. 395–430, 1989.

[11] R. K. Edwards, K. R. Kellner, C. L. Sistrom, and E. J. Magyari, "Medical student self-assessment of performance on an obstetrics and gynecology clerkship," *American Journal of Obstetrics & Gynecology*, vol. 188, no. 4, pp. 1078–1082, 2003.

[12] J. T. Fitzgerald, C. B. White, and L. D. Gruppen, "A longitudinal study of self-assessment accuracy," *Medical Education*, vol. 37, no. 7, pp. 645–649, 2003.

[13] D. A. Davis, P. E. Mazmanian, M. Fordis, R. Van Harrison, K. E. Thorpe, and L. Perrier, "Accuracy of physician self-assessment

compared with observed measures of competence: a systematic review," *The Journal of the American Medical Association*, vol. 296, no. 9, pp. 1094–1102, 2006.

[14] M. A. Mohammed and R. Holder, "Introducing analysis of means to medical statistics," *BMJ Quality & Safety*, vol. 21, no. 6, pp. 529–532, 2012.

[15] E. S. Escovitz, "Using senior students as clinical skills teaching assistants," *Academic Medicine*, vol. 65, no. 12, pp. 733–734, 1990.

[16] D. Nestel and J. Kidd, "Peer tutoring in patient-centred interviewing skills: experience of a project for first-year students," *Medical Teacher*, vol. 25, no. 4, pp. 398–403, 2003.

[17] M. Field, J. M. Burke, D. McAllister, and D. M. Lloyd, "Peer-assisted learning: a novel approach to clinical skills learning for medical students," *Medical Education*, vol. 41, no. 4, pp. 411–418, 2007.

[18] S. Buckley and J. Zamora, "Effects of participation in a cross year peer tutoring programme in clinical examination skills on volunteer tutors' skills and attitudes towards teachers and teaching," *BMC Medical Education*, vol. 7, article 20, 2007.

[19] S. A. Haist, J. F. Wilson, and N. L. Brigham, "Fourth-year students as teachers of the physical examination," *Academic Medicine*, vol. 71, no. 5, pp. 520–521, 1996.

[20] P. Weyrich, M. Schrauth, B. Kraus et al., "Undergraduate technical skills training guided by student tutors—analysis of tutors' attitudes, tutees' acceptance and learning progress in an innovative teaching model," *BMC Medical Education*, vol. 8, article 18, 2008.

[21] D. L. Kirkpatrick and J. D. Kirkpatrick, *Evaluating Training Programs*, Berrett-Koehler Publishers, 1994.

[22] V. Langendyk, "Not knowing that they do not know: Self-assessment accuracy of third-year medical students," *Medical Education*, vol. 40, no. 2, pp. 173–179, 2006.

Autism
Research and Treatment

Economics
Research International

Journal of
Biomedical Education

Nursing
Research and Practice

Child Development
Research

Journal of
Criminology

Journal of
Archaeology

International Journal of
Population Research

Education
Research International

Hindawi

Submit your manuscripts at
http://www.hindawi.com

Sleep Disorders

Depression Research
and Treatment

Journal of
Anthropology

Journal of
Addiction

Schizophrenia
Research and Treatment

Current Gerontology
& Geriatrics Research

Geography
Journal

Journal of
Aging Research

Psychiatry
Journal

Urban Studies
Research