Using Celebrities and Advertising Campaigns to Enhance Learning of Critical Review and Experimental Design, within an Inquiry-Oriented Biomedical Curriculum

Kerry Ann Dickson[a], and Bruce Warren Stephens[b]

Corresponding author: Kerry.Dickson@vu.edu.au
[a]College of Health and Biomedicine, Victoria University, Melbourne VIC 8001, Australia
[b]Department of Econometrics and Business Statistics, Monash University, Melbourne VIC 3145, Australia

Keywords: biomedical, popular culture, inquiry-oriented, scientific literacy, critical analysis

International Journal of Innovation in Science and Mathematics Education, *23*(6), 16-33, 2015.

Abstract

A challenge in learning to become a scientist is gaining skills in critical review and experimental design. Our aim was to measure the effectiveness of an inquiry-oriented learning (IOL) workshop which used popular culture and pseudoscience as stimuli for engagement. The workshop on critical review of scientific literature and best-practice experimental design consisted of a Socratic-seminar (i.e., collaborative, intellectual dialogue facilitated with open-ended questions) and a poster-defence. Students analysed the scientific publication which led Dr Oz (medical doctor and TV host) to falsely claim green coffee bean extract as a 'miracle' cure for obesity. Students also designed an experiment to test the effectiveness of an advertising campaign (e.g., Old Spice: 'The Man Your Man Could Smell Like') and presented their design in poster format. Students were assessed before and after the workshop. Post-test scores were higher than pre-test scores ($51.8 \pm 3.8\%$ vs $38.2 \pm 3.1\%$, $n = 25$, $p < 0.0001$) and were correlated ($p < 0.001$) with students' assignment marks. Students agreed that the workshop developed their ability to critically review scientific literature (79%) and to design experiments (63%). Our findings suggest that an IOL workshop, using popular culture and pseudoscience, improves skills in critical review and experimental design.

Introduction

Competence in critically reviewing original articles and in experimental design are important skills for scientists. Without them, pseudoscience flourishes. Based on nearly 10,000 responses, American undergraduate students' scientific literacy was shown to be only marginally higher than that of the American general public, with up to 40% of students believing in pseudoscientific theories (Impey, Buxner, Antonellis, Johnson and King 2011). Teaching scientific literacy has been shown to improve students' confidence in their ability to distinguish valid science from pseudoscience, their proficiency to summarise primary literature for a lay audience, their skills in critical review and experimental design, and their understanding of epistemological beliefs (Brownell, Price and Steinmam 2013; Bugarcic, Zimbardi, Macaranas and Thorn 2012; Duncan and Arthurs 2012; Gottesman and Hoskins 2013; Hoskins, Stevens and Nehm 2007; Murray 2014; Robertson 2012; Wenk and Tronsky 2011).

Although Faculty agree on the importance of students developing skills in scientific literacy, the definition and measurement of scientific literacy are problematic (Gormally, Brickman and Lutz 2012). The Organisation for Economic Co-operation and Development (OECD) definition is *'the capacity to use scientific knowledge, to identify questions and to draw* *evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity.'* A scientifically-literate student should be capable of evaluating the quality of scientific information on the basis of its source, research design, methods, quantitative data analysis and validity of interpretation (Gormally et al. 2012). Thus, scientific literacy is, basically, the knowledge and understanding of scientific concepts and evidence-based processes which are required for productive engagement in society - without resort to pseudoscience. Critical review of literature and knowledge of experimental design are integral to scientific literacy.

Research (Schroeder, Scott, Tolson, Huang and Lee 2007) which tabulates the *effect sizes* of various overlapping teaching strategies in science ranked the four most effective as: enhanced context (e.g., real-world linkages - such as use of popular culture), collaborative (e.g., open discussion - as in a Socratic seminar), questioning (e.g., high-order thinking questions - as in a poster-defence) and inquiry (e.g., inquiry-oriented learning).

Higher education students are attuned to multi-media (e.g., websites, television, smart phones) and popular culture (e.g., celebrities and advertising campaigns). Because students can relate to and are engaged with popular culture, educators have incorporated media (e.g., films, prime-time animation) into classes. From meta-analysis, the largest *effect size* (1.48) was found when teaching strategies incorporated examples to which students can relate (Schroeder et al. 2007). Popular media have been used to engage students in a range of disciplines including geography (Gill 2011), biology (Gardner, Jones and Ferzli 2009; Madhuri and Broussard 2008) health (Curch 2010; Diez, Pleban and Wood 2005) and information literacy (Ashton 2008; Peterson 2010; Tewell 2014). In general, these authors found that student engagement and performance improved.

A Socratic-seminar is collaborative, intellectual dialogue aimed at stimulating deep understanding of the ideas, values, arguments and evidence presented. The teacher acts as a facilitator, providing open-ended questions prior to the seminar and ensuring that the exchange of ideas during the seminar occurs in an open, rigorous, thoughtful manner. As students are responsible for the quality of the discussion, preparation by them is key. Socratic-seminars have been used to improve student engagement with analysing scientific articles on disease processes (Tangalakis, Hughes, Brown and Dickson 2014).

A poster-defence is a common form of presentation at scientific conferences. It visually summarises complex information, often using tables and graphs. In only a few minutes, the information is presented orally in a concise, logical and clear sequence - engaging the audience and encouraging them to ask challenging questions. Poster presentations have been used as an assessment tool, with students responding positively to the collaborative experience (Logan, Quiñones and Sunderland 2015).

Inquiry-oriented learning (IOL) is active, student-centred learning which is facilitated by a teacher. It engages by challenging students to think innovatively, methodically collect data, analyse with appropriate statistical tests, and interpret and explain the evidence. Critics (Kirschner, Sweller and Clark 2006) argue that, in order to learn, students require more structure than IOL offers. However, there is a general consensus, based on meta-analyses, that IOL is superior to traditional instruction - with an *effect size* of 0.5 to 0.65 (Furtak, Seidel, Iverson and Briggs 2012; Schroeder et al. 2007).

IOL has different levels of inquiry (Banchi and Bell 2008). With *simple-confirmation* IOL, students are asked to follow a detailed procedure - with predetermined results. With the next level, *structured* IOL, students conduct an experiment, collect data and explain their findings through evaluation and analysis. With *guided* IOL, students are given a scientific question, but are required to design an experiment, analyse the data and present their findings. Finally, with *open* IOL, students formulate their own research question.

Although *open* IOL fosters higher-order thinking, it is challenging and time-consuming for teachers to design and implement. It is also difficult for students to attain if it is not scaffolded throughout their educational development (Meyer, Meyer, Nabb, Connell and Avery 2013). Several authors (Gliddon and Rosengren 2012; Kazempour, Amirshokoohi and Harwood 2012; Loveys, Kaiser, McDonald, Kravchuk, Gilliham, Tyerman and Able 2014; Zimbardi, Bugarcic, Colthorpe, Good and Lluka 2013) have recently introduced an *open*, or near *open*, IOL research project into their classes and found that, in general, students perceived an improvement in their critical review, problem solving, data interpretation and communication skills.

The aim of the present pilot case-study was to measure the effectiveness of learning critical review of scientific literature and best-practice experimental design in a *guided* IOL workshop. The workshop, consisting of a Socratic-seminar and a poster-defence, used celebrities and advertising campaigns as stimuli for engagement. Thus, the novel aspect of our study is that we amalgamated popular culture and pseudoscience in a *guided* IOL workshop utilising four very effective teaching strategies.

Methods

This study was conducted at an institution with a high percentage of disadvantaged students, e.g., low socioeconomic status (SES), first in the family to attend university, English as a second language, working part- or full-time, or under-achieving.

Learning to critically review the primary literature, including understanding best-practice principles of experimental design, is scaffolded throughout the degree in Biomedical Science. In a year 1 unit called Foundations in Biomedical Science, students learn academic literacies, including critical review of lay and scientific articles. They start with the basics, such as the information conveyed by a graph and identification of the key claims. Then, in year 2, students collect and analyse physiological data and present their findings. In year 3, students are required to formulate their own research question.

Students in the present study were enrolled in a year 3 unit called Advanced Experimental Techniques. This unit aims to develop students' skills in independent thought, critical review, experimental design, data collection and interpretation of findings - while they learn a range of 'hands on' laboratory skills. Critical review of scientific literature and experimental design were revised at the start of semester. Then, students were exposed to a range of laboratory techniques, in the following order: animal handling, sterile surgery (operating on an anaesthetised rat), cell culture, tissue processing for light microscopy, immuno-histochemistry, measurement of gene expression using polymerase chain reaction, separation of proteins using polyacrylamide gel electrophoresis, bioinformatics and assessment of mitochondrial function. This 60 hour unit, which is recommended for students wishing to become medical researchers, was primarily taught in practical sessions via *guided* IOL. The within-semester assessment consisted of a grant proposal assignment (20%), a journal article assignment (20%), three laboratory reports (35%), and evaluation of laboratory competency (25%). The grant proposal assignment assessed student knowledge of critical review and experimental design.

IOL workshop

In week 1 of semester, students were given a 2 hour lecture on critical review and experimental design. The lecture revised the important points for critiquing a scientific article (Seals and Tanaka 2000), with particular emphasis on appropriate experimental design (e.g., sample size, longitudinal or cross-sectional design, minimising error). The IOL workshop, designed to be 3 hours in length, was conducted 10 days after the lecture on critical review and experimental design. The workshop consisted of two activities. The first was a Socratic-seminar, based on the scientific publication used by Dr Oz (medical doctor and TV host) to justify his claim that green coffee bean extract was a 'miracle' weight-loss supplement. The second activity was a poster-defence which utilised popular culture advertising campaigns. Lecture synopses and workshop instructions were provided on-line prior to the face-to-face classes. The IOL workshop was facilitated by a lecturer and observed by a sessional tutor.

Socratic-seminar

During the Socratic-seminar, students took turns at being an 'expert' who led the discussion, a 'contributor' who answered the expert's questions or a 'researcher' who provided back-up knowledge, via *Post-it* notes, to the 'contributor'. The instructions for the Socratic-seminar highlighted that peer reviewing of primary literature (evaluating the rationale, originality and significance of the hypothesis; the strengths and limitations of the experimental design; the quality of the data; the interpretation of the findings; and the directions for future studies) is integral to medical research. They also explained that obtaining skills in critical review is very challenging and time-consuming - requiring substantial practice. Furthermore, as scientific knowledge is influenced by social context and ethics, the instructions emphasised that science does not develop 'in a vacuum'. Finally, the instructions stressed the importance in the careers of medical researchers, health professionals and science writers of communicating with the public. For example, global issues, such as obesity or the spread of infectious diseases, will not be solved if pseudoscience thrives. Students were given a checklist for critical review of the primary literature, including the questions listed in Table 1.

Prior to the seminar, students were instructed to read several articles (Mubarak, Hodgson, Considine, Croft and Matthews 2013; Onakpoya, Terry and Ernst 2011) and to view some *YouTube* videos (*ABC News* 2012; Naughton 2010). Students were required to critically review the scientific article by Vinson, Burnham and Nagendran (2012), which Dr Oz used as evidence to claim that green coffee bean extract resulted in weight loss of 454 gm per week without a change to diet or exercise. This Vinson et al. (2012) article was published in a peer-reviewed journal (SJR: 0.68; Quartile 2 in 2013). The article was well-written and dealt with a topic of important physiological significance, given the health impacts of obesity. However, the article had several flaws, including that it was not double-blind, did not have a control for both doses of green coffee bean extract, was not a full cross-over study, had a short washout period, had a small *n* value, did not measure physical activity, lacked a breakdown of diet, used inappropriate statistical analysis, lacked standard error bars, did not account for side-effects of the coffee beans, and lacked an explanation for the weight reduction of subjects who received the placebo.

Table 1. Examples of critical review of the primary literature questions

What is the journal's Impact Factor?
Does the Title reflect the purpose, design and results of the study?
Is the Abstract a succinct and comprehensive summary?
Does the Introduction state what is known about the topic?
Is the study's sample appropriate (e.g., size, randomisation)?
Is there an appropriate control for each intervention?
Are there confounding variables?
Are the measurement techniques reliable, precise and valid?
Is the study design appropriate (e.g., double-blind, placebo-controlled)?
Is the statistical analysis appropriate?
Are the data clearly reported?
Are standard deviations/errors reported?
Are data presented for all of the important variables?
Are the conclusions supported by the data?
Are there other possible interpretations of the data?
Is the significance to the area of research reported fairly?
Are the limitations of the study adequately discussed?
Are there any conflicts of interest (e.g., financial)?

Poster-defence

During the class, student teams produced and displayed a scientific poster - electronically or on paper. Each team's poster was different and chosen from a list of popular culture marketing campaigns such as: *Old Spice - The Man Your Man Could Smell Like; Duracell lasts up to 30% longer than the competition; Colgate is different to ordinary fluoride toothpastes, it fights plaque bacteria for up to 12 hours; Olay Cleansing System delivers deep cleansing that is up to 4 times better than manual cleansing; Panadol - for pain relief you can start to feel in just 15 minutes.* Posters, which essentially consisted of information required for a grant application, were required to have a clear and informative Title; an Introduction; a concise, precise and testable aim/hypothesis; and a Methods section. The Methods section had to describe the independent variable(s) and how they were manipulated, the dependent variable(s) and how they were measured, factors that were kept constant, confounding factors, the *n* value, the procedural steps and the appropriate statistical analysis.

During the workshop, a student randomly chosen from each team was asked to orally present (2-3 minutes) the team's poster. This required the student to concisely summarise the main points and present the information in a logical and clear sequence. Remaining team members were allowed to add or clarify any information. During the workshop, each team critically reviewed each other team's poster - using *Post-it* notes. Students were asked to provide one positive comment, one negative comment and one question that reflected the reviewer's depth of understanding of experimental design. Students re-formed into their teams to discuss the peer review feedback, with time allocated for rebuttal. In this way, students received feedback - similar to a researcher receiving feedback from a grant application or a journal.

The instructions for the poster-defence highlighted that best-practice experimental design is a requirement of high-quality grant applications and gave links to supporting internet sites, e.g., www.3rs-reduction.co.uk, www.badscience.net and www.avid.org. Furthermore, the instructions gave an example of an experimental design to test a claim about an advertised household product, e.g., '*detergent X makes clothes whiter*'. By cleaning clothes in plain water and several types of detergent (including detergent X), one can compare the 'whiteness' of the outcomes. Students were instructed that they needed to consider how 'whiteness' can be measured, what factors (e.g., water temperature) may influence the outcomes, how many times the experiment must be repeated, etc.

Pre- and Post-Tests

All 26 students enrolled in the unit attended the lectures and the IOL workshop, and completed the voluntary pre- and post-tests. However, due to a missing page, the test data from one student was excluded. Students were unaware that they were going to complete pre- and post-tests, neither did these tests contribute to students' assessment in the unit. To evaluate the IOL workshop's impact on learning, the same test was given immediately before and immediately after the workshop. Using the lower- to higher-level thinking criteria of Anderson and Krathwohl (2001), an independent academic classified the test questions (see Appendix) as level 1 (remember, understand), level 2 (apply, analyse), or level 3 (evaluate, create). Respectively, there were 18, 12 and 6 test questions, marked out of 18, 12 and 12. Each student received a percentage score for each of the three levels of test question (e.g., level 1 had 18 questions worth 1 mark each: maximum 100%). The mean of these three percentages gave a student's test score (%), the levels being weighted equally in order to account for the disparity in numbers of questions and their marks. Whereas level 1 and 2 questions related to general aspects of experimental design, all level 3 questions related to the Vinson et al. (2012) article (Table 2).

Table 2. Illustrations of Level 1-3 critical review and experimental design questions

Level 1	Level 2	Level 3
Name a data base	Contrast independent and dependent variables.	Evaluate the appropriateness of the experimental approach (e.g., subjects, study design, key methods, data analysis).
List features of a graph	Contrast continuous and discrete variables.	Appraise the reasonableness of the conclusions.

Students were prevented from using any reference material provided by the lecturer (e.g., lecture synopses) but were allowed to use their own notes from reviewing the Vinson et al. (2012) article. No time limit was set for the tests, but the vast majority of students finished each within 20 minutes. Thus, the tests were open-ended and were open-book for level 3 questions only. The rationale here was that critically reviewing a scientific paper is time-consuming and details are difficult to remember without notes. Pre- to post-test difference in scores (n = 25) measured learning which took place during the IOL workshop.

Survey

The authors of the present study constructed a survey (see Appendix 1) to determine the level of student satisfaction with the IOL workshop. Of the 13 survey items, four were negatively-keyed. The nine positively-keyed items surveyed the following: ability to critically review, ability to design experiments, improvement in knowledge base, perceived interactivity, understanding of the teacher's instructions, understanding of workshop content, perceived level of challenge, awareness of the difficulty in understanding scientific content and usefulness of the workshop. Examples of the form of the items were *'It developed my ability to critically review a scientific paper.'* and *'Compared with the normal tutorial, I found it more interactive.'* Twenty-four students completed the voluntary and anonymous survey, which consisted of 13 items scored on a five-point Likert scale (strongly agree = 5, agree = 4, neither agree nor disagree = 3, disagree = 2, strongly disagree = 1).

Grant Proposal Assignment

For unit assessment purposes, learning of critical review and experimental design was measured by marks on a grant proposal assignment, due four weeks after the IOL workshop. Students devised a novel hypothesis (different from any in the medical research literature or devised by other students), completed a review - including critical review of pivotal primary literature - and designed an experiment to solve the problem. This *open* IOL task required students to think creatively and critically. Although pre- and post-test scores were obtained for 25 students, 2 of these did not submit a grant proposal assignment.

Student Demographics

Students' age, sex, grade point average (GPA) and home address zip code were obtained from institutional records. SES was determined, via zip code, from the national Index of Disadvantage. This study received approval from the institution's Human Research Ethics Committee (HREC).

Statistical Analysis

Scores on the pre- and post-tests ($n = 25$) for each of the three levels of question were compared using a repeated measures, two-way ANOVA with Tukey-Kramer *post hoc* comparisons. Wilcoxon's signed-rank procedure was used to test whether the median scores ($n = 24$) for the positively-keyed survey questions exceeded 3 (Likert: neither agree nor disagree). Pearson's correlation coefficient was calculated for the relationship between pre-test score and grant proposal assignment mark, and for the relationship between post-test score and grant proposal assignment mark. The relationship of post-test score and student demographic variables (age, sex, SES, GPA) to grant proposal assignment mark was analysed via multiple regression analysis ($n = 23$). All statistical tests were conducted in *SPSS version 22*, except for the Tukey-Kramer *post hoc* comparisons (*Excel*). Results were expressed as mean ± SEM or as %.

Results

Of the 26 students enrolled in the unit, 14 were male; and 31% were low, 50% medium and 19% high SES. The two-way ANOVA showed significant ($p < 0.0001$) main effects for test (pre- or post-), for level of question and for the interaction between test and level. Post-test scores were higher ($p < 0.0001$) than pre-test scores (Figure 1). Test scores decreased ($p < 0.0001$) as question level increased. *Post hoc* analysis showed that pre- and post-test scores were not significantly different at levels 1 and 2, but were at level 3 ($p < 0.01$).

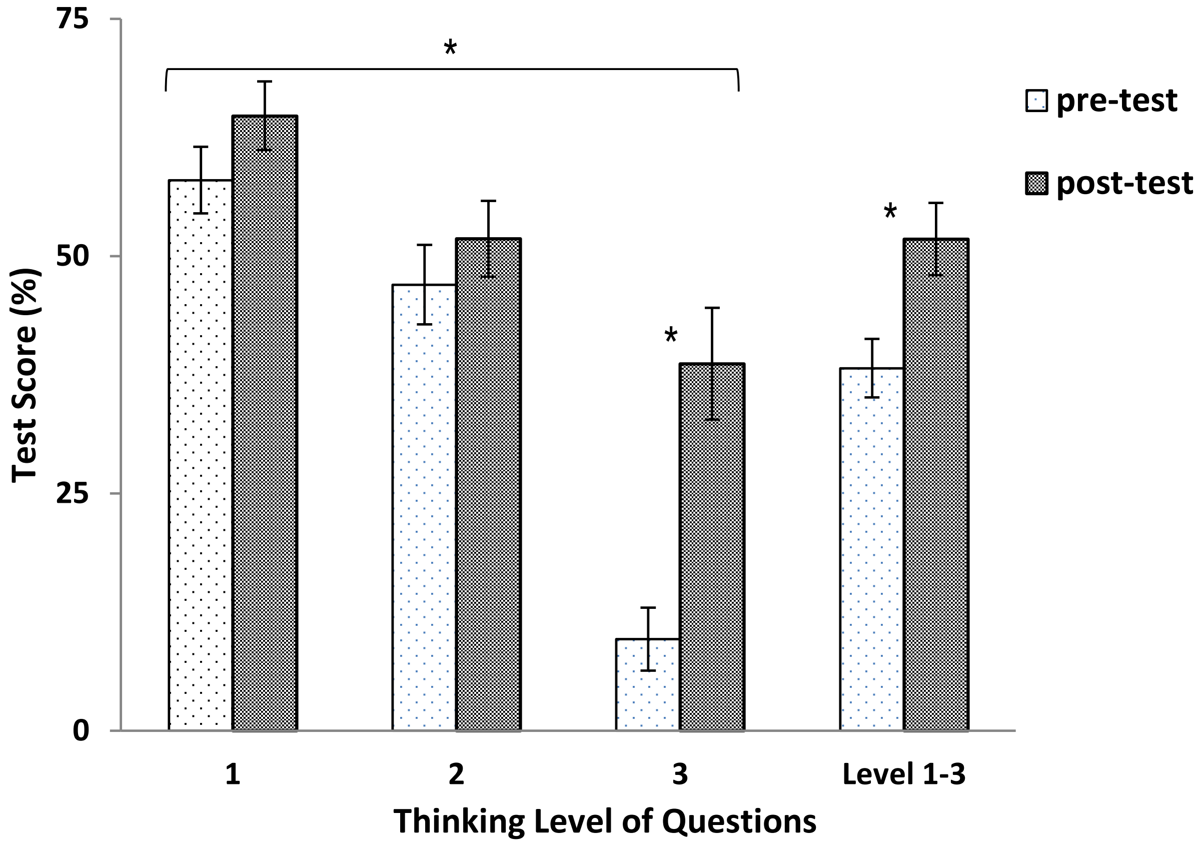

Figure 1. Pre- and post-test scores (mean ± SEM, n = 25): for questions at each level of thinking (1: remember, understand; 2: apply, analyse; 3: evaluate, create) and for the test (level 1-3 combined). Post-test scores were higher (p < 0.0001) than pre-test scores. Test scores decreased (p < 0.0001) as thinking level increased. Pre- to post-test improvement in score was greater (p < 0.01) for higher-order thinking (level 3). Significant differences are indicated by *.

Figure 2 shows the percentage of students who strongly agreed or agreed with the survey items. All items had median Likert scale scores significantly ($p < 0.05$) greater than 3, except for students' perceived level of challenge. In the open-ended section of the survey, 58% of students stated that they enjoyed the interactive, engaging, and collaborative aspect; 17% stated that it was too challenging; and 29% commented on the improvement in their ability to critically review papers.

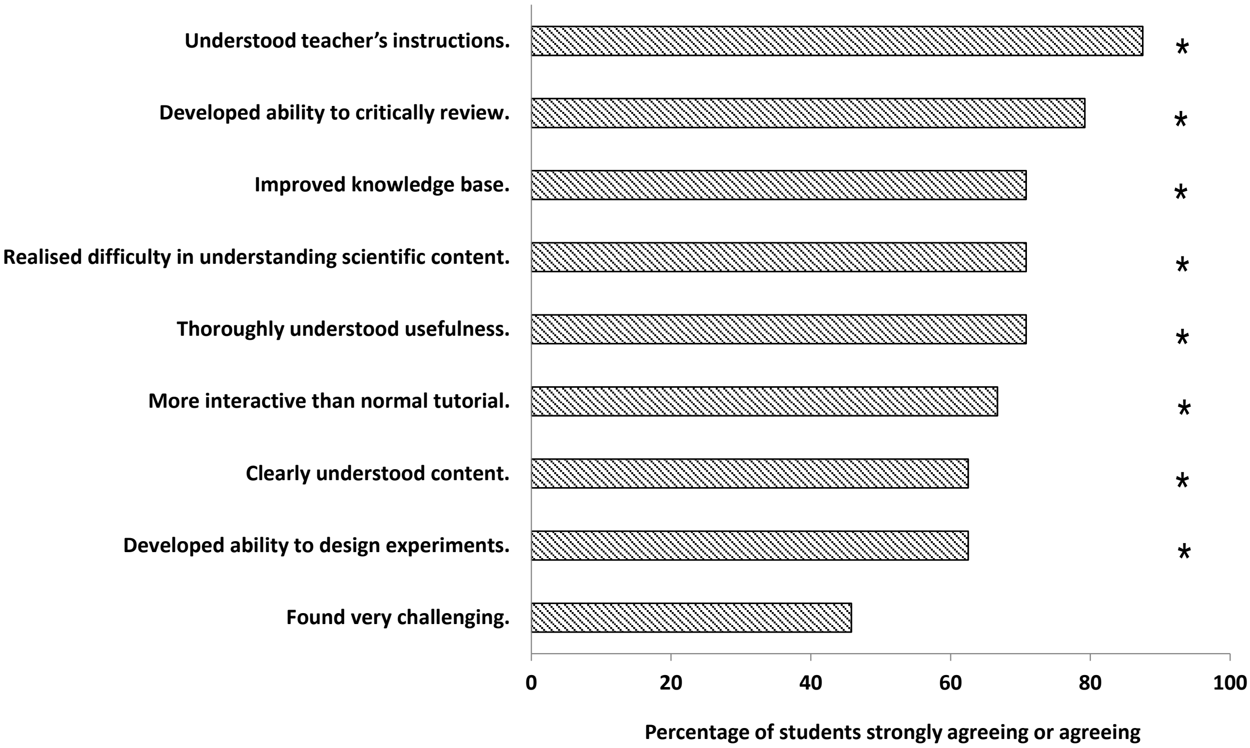

Figure 2. Student perceptions of the IOL workshop (percentage of students strongly agreeing or agreeing, n = 24). Survey items with a median Likert scale score significantly ($p < 0.05$) greater than 3 (neither agree nor disagree) are indicated by *.

Four weeks after the IOL workshop, students submitted a grant proposal assignment. The marks (%) for this assessment did not differ significantly with sex, SES or GPA. However, they were positively related to age ($p < 0.028$) and post-test score ($p < 0.001$): *Grant proposal mark = -11.04 + 2.55Age + 0.33Post-test score* ($r^2 = 0.5666$, n = 23). Although the grant proposal mark was positively related to the pre-test ($r = 0.6014$, $p < 0.01$, n = 23) and post-test ($r = 0.6675$, $p < 0.0001$, n = 23, Figure 3) scores, Pearson's correlation coefficient was higher for the post-test.

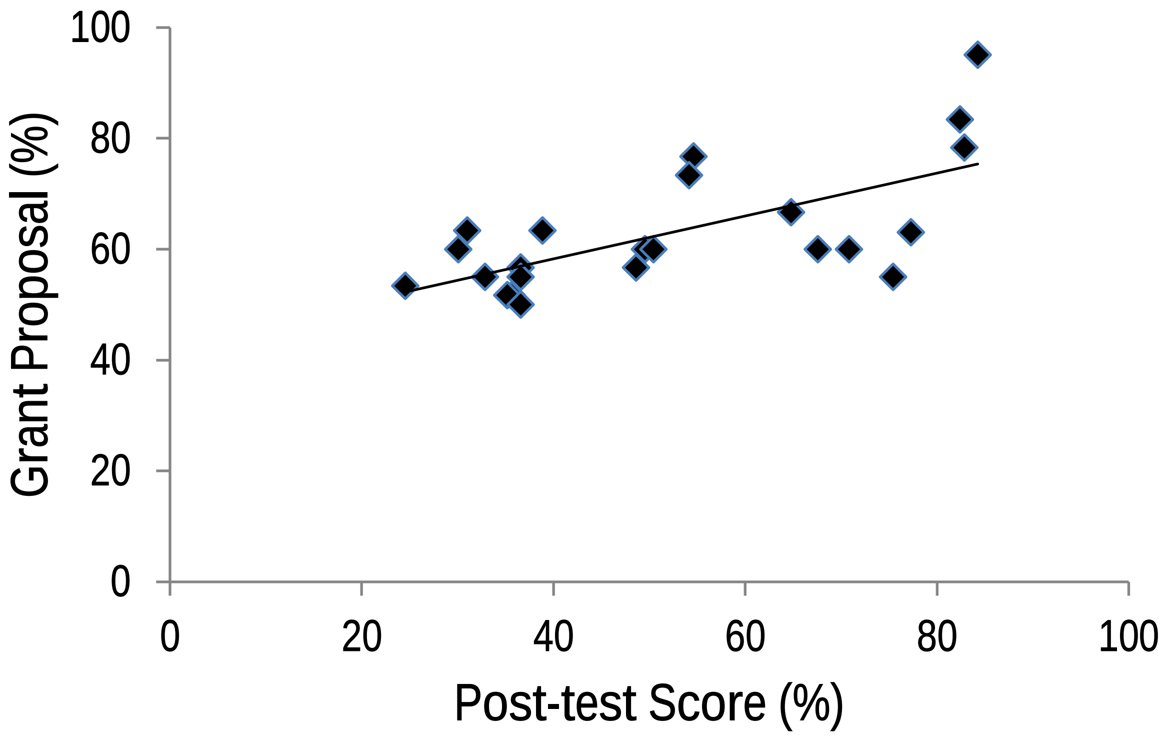

Figure 3. Relationship of post-test score (%) to student mark (%) for the grant proposal assignment (r = 0.6675, p < 0.001, n = 23).

Discussion

Our study is the first to amalgamate popular culture (e.g., celebrities and advertising campaigns) with a Socratic-seminar and poster-defence in a *guided* IOL workshop. Based on differences in pre- and post-test scores, we found that our IOL workshop using popular culture as a stimulus for engagement improved student performance in critically reviewing primary literature and in understanding best-practice experimental design. The improvement occurred primarily in level 3 questions (evaluate, create), which suggests that our IOL workshop stimulated higher-order thinking. Students perceived that the workshop was effective in improving their ability to critically review literature and to design experiments. Students who scored highest on the post-test tended to score highest on the grant proposal assignment.

However, our study has several limitations. It was carried out on a small number of participants, in only one unit, at a single institution - limiting the generalisation of the findings. As we did not compare our workshop with and without popular culture, our study needs to be repeated with an improved experimental design. As the same questions were used in the pre- and post-tests, and they were completed in a single day, a test-effect may have contributed to changes in test scores. Our study needs to be repeated without this confounding factor. Neither the lecturer nor the students were blind to the intervention. The lecturer's quality of delivery and students' preference for interactive, fun classes may have biased the results. Hence, our interpretation is very preliminary and requires affirmation by further work. Our survey was self-constructed and rudimentary. Future work should use a previously published and validated survey. Our approach may not be easily transferable to other disciplines. Finally, we did not account for other potentially-relevant factors including English language skills, metacognition, motivation, and sense of responsibility for learning.

Our results suggest that students benefited from our workshop in terms of higher-order rather than lower-order thinking (Anderson and Krathwohl 2001). In the pre-test, students performed well on the level 1 and 2 questions. For example, around 75% of students correctly defined a double-blind clinical trial (level 1) and correctly explained why scientists incorporate a control condition such as a placebo (level 2). However, only 60% of students attempted any level 3 questions - suggesting that the remainder either did not read or did not understand the Vinson et al. (2012) paper. Of the students who attempted the pre-test level 3 questions, very few were able to synthesise the knowledge from their level 1 and 2 answers and relate it to the paper by Vinson et al. (2012). For example, most students failed to mention the paper's lack of double-blinding or its inappropriate placebo.

During the Socratic-seminar, discussion focused on the bias inherent in non-blind and single-blind trials and listed the requirements of a placebo (e.g., same size, colour, taste, dose, number of tablets/day). Discussion of the Vinson et al. (2012) paper revealed that the low-dose of green coffee bean extract was taken twice daily while the placebo was taken three times daily. Although Vinson et al. (2012) acknowledged this limitation in their paper, students suggested that the paper's title was misleading in its use of 'double-blind' and 'placebo-controlled'. Thus, in our level 3 post-test questions, students mentioned the inappropriate placebo or the lack of blinding in the paper. Many students commented on their increased confidence to critically review a paper, their surprise that papers with numerous flaws were published in high-quality journals and their greater understanding of the machinations of scientific and pseudoscientific publishing.

Students perceived that the workshop developed their ability to critically review literature (Likert 4.00 ± 0.22) and to design experiments (3.58 ± 0.20), improved their knowledge base (3.71 ± 0.20) and was useful (3.96 ± 0.23). These results suggest that students considered the workshop worthwhile. Our IOL workshop ran over time (4.5 hours in total). The sessional tutor who observed the workshop concluded that most students were highly engaged in the IOL activities and that this may explain the elongated duration. However, the reason may have been that the workshop had flaws in design or instructions. In our study, both pre- and post-test scores were disappointingly low, particularly pre-test scores. The probable reason was that many students did not adequately prepare for the workshop – perhaps because it did not form part of the formal assessment in the unit.

Marks (%) for the grant proposal assignment were higher than the post-test scores (%). One reason may be the different criteria for marking. Unlike the post-test, the grant proposal marking rubric included criteria on quality of literature reviewed; innovative thought; experimental design and analysis; and referencing and writing style. Another reason may have been revealed by the survey, which showed that the IOL workshop made 71% of students aware of the difficulty of critical review and experimental design. This awareness may have stimulated students to work more conscientiously on their grant proposal. Students' marks for the grant proposal were positively related to test scores, in particular to post-test scores. Our results suggest that students who were receptive to the content of the lecture and IOL workshop were more likely to perform well in this assignment. Interestingly, GPA was not significantly correlated with grant proposal mark.

In the academic year subsequent to that of the present study, we improved our IOL workshop (Socratic-seminar and poster-defence). This was done in consultation with the tutor who had observed the original workshop and documented its most, and least, engaging aspects. The improvements included making participation in the workshop part of the student's competency mark, requiring student facilitation of the workshop, providing instructions that were more explicit and detailed, and having the lecturer organise the student teams in advance. These changes increased the quality of the IOL workshop without it running over time.

Our Socratic-seminar was improved by requiring each student team to prepare a different section of the Vinson et al. (2012) paper (e.g., Introduction) prior to the workshop and by requiring each team to facilitate the seminar for their section. Thus, each facilitating team had to be very well-prepared, keep the topic focussed, coax reluctant participants, limit contributions from dominant students, and encourage students to justify their contributions with evidence. Meanwhile, the remainder of the class had to be well-prepared, forward their opinions and listen respectfully. Our poster-defence was improved by requiring student teams to create their posters prior to the workshop, to limit their critical review to one insightful comment per poster and to work collaboratively during their poster-defence rebuttal.

After we implemented the above improvements, the proportion of students who strongly agreed or agreed that the workshop developed their ability to critically review increased from 79% to 100%; developed their ability to design experiments, from 63% to 90%; improved their knowledge base, from 71% to 90%; was useful, from 71% to 95%. However, compared with previous semesters (63.3 ± 1.3, $n = 9$) in which the content was conveyed only via lectures, mean semester mark for the grant proposal assignment was unchanged in the semesters which had the IOL workshop (63.4 ± 0.2, $n = 2$).

IOL research projects have been introduced into practicals in toxicology (Gliddon and Rosengren 2012) and plant biology (Loveys et al. 2014). Similarly, Zimbardi et al. (2013) found that introducing IOL to year 1 and 2 undergraduates improved learning of scientific method (e.g., incorporation of dependent and independent variables into hypothesis writing and of appropriate statistical information in figure legends). Qualitative assessment of student perceptions showed that IOL was a positive experience and that students perceived an improvement in their understanding of scientific inquiry (Kazempour et al. 2012). Collins and Calhoun (2014) asked year 1 undergraduates to rate the laboratory sessions in their integrated science unit, finding that the favourite was the one in which students designed and implemented their own experiments. Our findings also suggest that IOL was perceived as a positive experience. Student evaluations of our unit and its teaching staff were high. For example, following the IOL workshop and grant proposal assignment, an unsolicited student email stated, *'You have provided an excellent challenge in this assignment. Lots of thinking involved.'*

Surveys reveal that student perceptions of their ability to write reports, formulate questions, devise experiments and deliver oral presentations improved in response to IOL (Loveys et al. 2014). However, these authors also noted that students struggled with developing a hypothesis, statistical analysis, and interpretation of data - even though their students had completed a statistics unit. In response, Loveys et al. (2014) provided on-line tutorials as well as consultations with a biometrician. We agree that students struggle to develop an innovative, testable hypothesis and to design an appropriate experiment. Anecdotally, we note that students who perform well on the grant proposal assignment are passionate and well-organised - resulting not only in a novel aim but a grant which is focused, precise, concise, logical and conforms to the conventions of scientific presentation. However, although other students may be enthusiastic, resulting in a novel aim, they lack skills in in-depth analysis. Thus, clearly, we need to provide more scaffolding opportunities in years 1 and 2 to support these students.

Embedding learning about the nature of science into astronomy classes has been shown to improve student self-reported confidence in their ability to distinguish valid science from pseudoscience (Duncan and Arthurs 2012). Year 1 undergraduates' ability to read scientific literature has been shown to improve slightly if students are given explicit instructions (Wenk and Tronsky 2011). Providing year 1 undergraduates with a unit on scientific thinking has been shown to improve their skills in critical review and experimental design, their understanding of scientists as people and their own epistemological beliefs (Gottesman and Hoskins 2013). Hoskins et al. (2007) showed similar findings when using their teaching strategy with later-year undergraduates, the majority of whom were educationally disadvantaged. More recent work has shown that skills in critical review and experimental design can be improved via short duration, journal-club workshops (Robertson 2012) and practical modules (Bugarcic et al. 2012). Brownell et al. (2013) found that an intensive writing course in neuro-immunology improved student perceptions of their ability to read primary articles and summarise the content for a lay audience. Our findings confirm earlier work showing that a range of teaching strategies can be implemented to increase students' confidence to critically review scientific papers and to distinguish science from pseudoscience.

Murray (2014) used Process Oriented Guided Inquiry Learning (POGIL) activities to teach critical analysis. Students in her biochemistry units reviewed 1-3 research articles - chosen because they were believed to be of interest to students, were well-written and were from well-regarded journals. Murray (2014) found that, when both assessments used similar articles and questions, examination scores (76.1 ± 16.1%, mean ± SD) were higher than pre-test scores (26.3 ± 11.3%). Murray (2014) also showed that most students considered that the POGIL activities improved their skills in critical reading (Likert score: 3.5) and their confidence in designing biochemical experiments (3.1). In general, our results support Murray's conclusions. Our Likert scores for developing ability to critically review (4.0) and ability to design experiments (3.6) were higher than Murray's (2014), particularly in the subsequent year when we had implemented improvements to our IOL workshop (4.7 and 4.5, respectively). Our higher Likert scores may be because we chose an article more recent and controversial than those chosen by Murray (2014). We similarly recorded low pre-test scores (38.2 ± 3.1%) and, although our post-test scores were lower than Murray's (2014), her experimental period was substantially longer - a full semester.

One of the greatest challenges for educators is engaging students and motivating them to learn, particularly if the cohort is diverse in aptitude, metacognitive skills, non-academic responsibilities and financial constraints. One way to stimulate engagement is to incorporate popular culture. Media issues (e.g., childhood obesity) and films (Diez et al. 2005; Gardner et al. 2009) have been proposed to encourage biology and health teachers to facilitate students' exploration of the cultural, moral and religious undertones of scientific research. Using IOL in practical classes (e.g., embryonic development) and popular culture in lectures (e.g., lay books on pregnancy), Madhuri and Broussard (2008) redesigned their developmental biology unit. Compared with their conventional classes, they found an improvement in student engagement and performance.

Using prime-time animation programmes (e.g., *The Simpsons*), Curch (2010) found that students perceived classes on ageing to be interesting, fun and engaging - making the content more easily understood. Among other activities, Gill (2011) lectured biogeography using a map of *Middle Earth* and video clips from *Lord of the Rings*. He found that students perceived the teaching approach helpful but not more so than conventional lectures.

Peterson (2010) used a video clip from *The School of Rock* to illustrate collaborative brainstorming, gathering information from multiple resources, planning a project and creating an original product - an approach that essentially mirrors academic research. Using pre- and post-tests, Tewell (2014) examined the effectiveness of a lesson on information literacy, with and without using television comedies. Unlike his control condition, he found that pre- to post-test scores improved significantly (22%) with the comedies. Although his statistical analysis had some deficiencies (e.g., a negative standard deviation value), he found that the comedies were enjoyable and focused students' attention, particularly on resource evaluation (e.g., the authority of a website). Our finding that pre- to post-test scores improved 35% with our IOL workshop utilising television celebrities and advertising campaigns supports Tewell's findings.

Using a popular American tabloid magazine which printed blue dots that were claimed to be empowered by psychics, Ashton (2008) taught scientific literacy to year 1 psychology students. They were required to conduct and write-up an experiment testing the efficacy of the *psychic dots*. Student perceptions of the teaching approach showed that they considered that their knowledge of hypothesis testing, independent variables and dependent variables had improved (Likert scores, rescaled from his 7-point to a 5-point scale: 3.7 and 3.8). Although students found Ashton's (2008) approach interesting (3.5), they did not consider it worthwhile (3.3). In contrast to Ashton (2008), we found that students thoroughly understood the usefulness (4.0) of the teaching approach, particularly after we had made improvements to our workshop (4.4).

There were several advantages to our using the Dr Oz and the advertising campaign video clips, including that they were short, topical, comedic, parodic and presented social commentary. We conducted our IOL workshop for the first time in 2014. Interestingly, later in that year, Dr Oz was questioned by a US Senate committee (McCaskill 2014) regarding his promotion of bogus weight-loss supplements, including green coffee bean extract. Subsequently, the Vinson et al. (2012) publication was retracted - with the notification, *'The sponsors of the study cannot assure the validity of the data.'* When we ran our IOL workshop again in 2015, we were pleased to see the level of student engagement - evidenced by the fact that all students had discovered for themselves the Senate hearings and the retraction.

Based primarily on the taxonomy of Bloom and his co-workers, learning has been categorised (Anderson and Krathwohl 2001) into domains: *cognitive* (knowledge, thinking), *affective* (attitude, emotion) and *psychomotor* (skills, doing). Our students' comments indicated that they enjoyed learning which incorporated popular culture. During the workshop, numerous students were amazed that a medically-qualified person (Dr Oz) would endorse a commercial product without fully investigating its claims of efficacy. Furthermore, reference to humorous advertising campaigns appeared to arouse positive emotions, such as curiosity, amusement and collegiality. Thus, it is possible that the improvement we found in student learning was facilitated through the affective domain. Although there are different theories of cognition, remembering is stronger when the stimulus captures attention. It is possible that popular culture acts as a gateway to learning, stimulating interest and improving recall of course material.

In conclusion, our novel approach to *guided* IOL, which amalgamated popular culture and pseudoscience, provided a case-study which students found engaging and effective for the learning of critical review of scientific literature and best-practice experimental design. Future work may entail a simulation which involves cross-year peer collaboration (Sloman and Thompson 2010). During our IOL workshop, year 3 students could mimic journal editors or reviewers, year 2 students could mimic scientific authors or medical TV hosts and year 1 students could mimic members of the general public who are interested in weight control.

References

ABC News: Weight loss, metabolism and green coffee bean extracts Dr Ken and Dr Oz. (2012). Retrieved July 15, 2015, from http://www.youtube.com/watch?v=KoYgpuG87lQ.

Anderson, L.W., & Krathwohl, D.R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. (1st ed). New York, NY: Longman.

Ashton, W. A. (2008). Using the psychic blue dot to teach about science and pseudoscience. *Journal of Instructional Psychology, 35*(4), 409–412.

Banchi, H., & Bell, R. (2008). The many levels of inquiry. *Science and Children, 46*(2), 26–29.

Brownell, S. E., Price. J. V., & Steinman, L. (2013). A writing-intensive course improves biology undergraduates' perception and confidence of their abilities to read scientific literature and communicate science. *Advances in Physiology Education, 37*, 70–79.

Bugarcic, A., Zimbardi, K., Macaranas, J., & Thorn, P. (2012). An inquiry-based practical for a large, foundation-level undergraduate laboratory that enhances student understanding of basic cellular concepts and scientific experimental design. *Biochemistry and Molecular Biology Education, 40*(3), 174–180.

Collins, E. S., & Calhoun, T. R. (2014). Raising the bar in freshman science education: Student lectures, scientific papers, and independent experiments. *Journal of College Science Teaching, 43*(4), 26–35.

Curch, L. M. (2010). Using prime-time animation to engage students in courses on aging. *Gerontology and Geriatrics Education, 31*, 361–382.

Diez, K. S., Pleban, F. T., & Wood, R. J. (2005). Lights, camera, action: Integrating popular film in the health classroom. *Journal of School Health, 75*(7), 271–275.

Duncan, D. K., & Arthurs, L. (2012). Improving student attitudes about learning science and student scientific reasoning skills [Electronic version]. *Astronomy Education Review, 11*(1), 2009067.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research, 82*(3), 300–329.

Gardner, G. E., Jones, M. G., & Ferzli, M. (2009). Popular media in the biology classroom: Viewing popular science sceptically. *The American Biology Teacher, 71*(6), 332–335.

Gill, R. (2011). Effective strategies for engaging students in large-lecture, non-majors science courses. *Journal of College Science Teaching, 41*(2), 14–21.

Gliddon, C. M., & Rosengren, R. J. (2012). A laboratory course for teaching laboratory techniques, experimental design, statistical analysis, and peer review process to undergraduate science students. *Biochemistry and Molecular Biology Education*, *40*(6), 364–371.

Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a test of scientific literacy skills (TOSLS): measuring undergraduates' evaluation of scientific information and arguments. *CBE-Life Sciences Education, 11*, 364–377.

Gottesman, A. J., & Hoskins, S. G. (2013). CREATE cornerstone: Introduction to scientific thinking, a new course for STEM-interested freshmen, demystifies scientific thinking through analysis of scientific literature, *CBE-Life Sciences Education, 12*, 59–72.

Hoskins, S. G., Stevens, L. M., & Nehm, R. H. (2007). Selective use of the primary literature transforms the classroom into a virtual laboratory. *Genetics, 176*, 1381–1389.

Impey, C., Buxner, S., Antonellis, J., Johnson, E., & King, C. (2011). A twenty-year survey of science literacy among college undergraduates. *Journal of College Science Teaching, 40*(4), 31–37.

Kazempour, M., Amirshokoohi, A., & Harwood, W. (2012). Exploring students' perceptions of science and inquiry in a reform-based undergraduate biology course. *Journal of College Science Teaching, 42(*2), 38–43.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*, 75–86.

Logan, J. L., Quiñones, R., & Sunderland, D. P. (2015). Poster presentations: Turning a lab of the week into a culminating experience. *Journal of Chemical Education, 92*, 96−101.

Loveys, B. R., Kaiser, B. N., McDonald, G., Kravchuk, O., Gilliham, M., Tyerman, S., & Able, A. J. (2014). The development of student research skills in second year plant biology. *International Journal of Innovation in Science and Mathematics Education, 22*(3), 15–25.

Madhuri, M., & Broussard, C. (2008). "Do I need to know this for the exam?" Using popular media, inquiry-based laboratories, and a community of scientific practice to motivate students to learn developmental biology. *CBE-Life Sciences Education, 7*, 36–44.

McCaskill, C. (2014, June 17). *Protecting consumers from false and deceptive advertising of weight-loss products*. Retrieved July 15, 2015, from http://www.commerce.senate.gov/public/index.cfm?p=Hearings.

Meyer, D. Z., Meyer, A. A., Nabb, K. A., Connell, M. G., & Avery, L. M. (2013). A theoretical and empirical

exploration of intrinsic problems in designing inquiry activities. *Research in Science Education, 43*, 57–76.
Mubarak, A., Hodgson, J. M., Considine, M. J., Croft, K. D., & Matthews, V. B. (2013). Supplementation of a high-fat diet with chlorogenic acid is associated with insulin resistance and hepatic lipid accumulation in mice. *Journal of Agriculture and Food Chemistry, 61*, 4371–4378.
Murray, T. A. (2014). Teaching students to read the primary literature using POGIL activities. *Biochemistry and Molecular Biology Education, 42*(2), 165–173.
Naughton, T. (2010, October 28). *Big Fat Fiasco*. Retrieved July 15, 2015, from https://www.youtube.com/watch?v=exi7O1li_wA&list=PLCD72F4109EDC4BD8&index=2.
Onakpoya, I., Terry, R., & Ernst, E. (2011). The use of green coffee extract as a weight loss supplement: A systematic review and meta-analysis of randomised clinical trials [Electronic version]. *Gastroenterology Research and Practice, 382852*.
Peterson, N. (2010). It came from Hollywood: Using popular media to enhance information literacy instruction. *College and Research Libraries News, 71*(2), 66–74.
Robertson, K. (2012). A journal club workshop that teaches undergraduates a systematic method for reading, interpreting, and presenting primary literature. *Journal of College Science Teaching, 41*(6), 25–31.
Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T. Y., & Lee, Y. H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching, 44*, 1436–1460.
Seals, D. R., & Tanaka, H. (2000). Manuscript peer review: A helpful checklist for students and novice referees. *Advances in Physiology Education, 23*(1), 52–58.
Sloman, K., & Thompson, R. (2010). An example of large-group drama and cross-year peer assessment for teaching science in higher education. *International Journal of Science Education, 32*(14), 1877–1893.
Tangalakis, K., Hughes, K., Brown, C., & Dickson, K. (2014). The use of explicit teaching strategies for academic staff and students in bioscience foundation subjects. *International Journal of Innovation in Science and Mathematics Education, 22*, 42–51.
Tewell, E. C. (2014). Tying Television comedies to information literacy: A mixed-methods investigation. *The Journal of Academic Librarianship, 40*, 134–141.
Vinson, J. A., Burnham, B. R., & Nagendran, M. V. (2012). Randomized, double-blind, placebo-controlled, linear dose, crossover study to evaluate the efficacy and safety of a green coffee bean extract in overweight subjects. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy, 5*, 21–27.
Wenk, L., & Tronsky, L. (2011). First-year students benefit from reading primary research articles. *Journal of College Science Teaching, 40*(4), 60–67.
Zimbardi, K., Bugarcic, A., Colthorpe, K., Good, J. P., & Lluka, L. J. (2013). A set of vertically integrated inquiry-based practical curricula that develop scientific thinking skills for large cohorts of undergraduate students. *Advances in Physiology Education, 37*, 303–315.

Appendix 1.

Level 1: remember, understand. **18 marks**

1. A journal has *blind peer-review*. State what this means.
2. Name an internet site which rates the quality of a scientific journal.
3. Name one *Australian Category 1* funding body.
4. State what is measured by an *h-index*.
5. Name a database that is used to search for scientific papers.
6. Describe an Abstract.
7. Define *null hypothesis*.
8. Explain what scientists mean by *experimental design*.
9. Define *control group*.
10. Define *random sample*.
11. Define *double-blind clinical trial*.
12. Define *cross-over study*.
13. Explain why the Methods are described in a way that they can be reproduced.
14. Give an example of *experimental error*.
15. Explain why statistical techniques are important to scientists.
16. State what the acronym *ANOVA* denotes.
17. List 4 important features of a graph.
18. State the section of a paper in which the interpretation of the findings should appear.

Level 2: apply, analyse. **12 marks**

1. Of the following list Ng K., Smith F., Bloggs J., explain the likely role of each author.
2. Explain why we avoid uncontrolled variables when designing an experiment.
3. Explain why randomisation is important to experimental design.
4. Explain how you know whether your sample size is large enough.
5. Contrast independent and dependent variables.
6. Contrast qualitative and quantitative variables.
7. Contrast continuous and discrete variables.
8. Contrast longitudinal and cross-sectional studies.
9. Contrast homogenous and heterogeneous groups.
10. Contrast accuracy and reliability.
11. Explain whether we should believe everything that is published in a well-regarded scientific journal.
12. Discuss whether Science is *subjective* or *objective*.

Level 3: evaluate, create. **12 marks**

The following questions relate to the paper by Vinson et al., (2012).

1. Explain whether the role of green coffee bean extract on weight control is an important experimental question.
2. Explain whether a clear and testable hypothesis is presented.
3. Evaluate the appropriateness of the experimental approach (e.g., subjects, study design, key methods, and data analysis).
4. Explain whether the results are properly presented and valid.
5. Appraise the reasonableness of the conclusions (e.g., appropriate interpretation of the data).
6. Evaluate the consequences of the paper’s findings (e.g., to the general public, scientific community, regulatory bodies).

Survey

1. It developed my ability to critically review a scientific paper.
2. It developed my ability to design an experiment.
3. It improved my knowledge base.
4. Compared with a “normal” tutorial, I found it more interactive.
5. It did NOT provide me with any useful new ways of studying material.
6. It did NOT contribute to my understanding of good experimental design.
7. I clearly understood the teacher’s instructions.
8. It did NOT improve my confidence to critical review scientific papers.
9. I clearly understood its content.
10. I found it very challenging.
11. It added NO more to my knowledge base than what I obtained from the lectures.
12. It helped me realise the difficulty in understanding scientific content.
13. I thoroughly understood its usefulness.