COMPUTER ASSISTED SELF AND PEER ASSESSMENT RATINGS (CASPAR): APPLICATIONS, CHALLENGES AND OPPORTUNITIES

Peter Lugosi, PhD
School of Services Management
Bournemouth University
Fern Barrow
BH12 5BB
plugosi@bournemouth.ac.uk
01202 965219

Biosketch
Peter Lugosi, PhD lectures in hospitality management at the School of Services Management, Bournemouth University. His previous work has considered the relationship between social and commercial forms of hospitality in contemporary society, research ethics and the relationship between hospitality and regeneration.

The data which this paper is based on was gathered through a project funded by the Higher Education Academy Network for Hospitality, Leisure, Sport and Tourism Pedagogic Research and Development Fund 2007/08 (Round 9).

COMPUTER ASSISTED SELF AND PEER ASSESSMENT RATINGS (CASPAR): APPLICATIONS, CHALLENGES AND OPPORTUNITIES

ABSTRACT

This paper discusses the application of and challenges associated with CASPAR (Computer Assisted Self and Peer Assessment Ratings) – a software tool designed to aid the administration of self and peer assessment (SPA). CASPAR was piloted on eight units from six subject areas (Hospitality, Leisure, Sport, Tourism, Retail and Events). Data was gathered through questionnaires from 146 students. The paper discusses the functionality of CASPAR as well as its application in learning, teaching and assessment. The paper also compares the SPA process and application of CASPAR in a hospitality operations management unit and other units to identify good practice.

Keywords: self assessment, peer assessment, computer aided assessment, group work, transferable employability skills
INTRODUCTION

Group work remains a fundamental part of teaching in hospitality, leisure, sport, tourism and events and it continues to present opportunities and challenges for students and teaching staff (Hassanien, 2006). Group work helps to develop key transferable employability skills (Boud, Cohen & Sampson, 1999; Macpherson, 1999), but it is often accompanied by tensions arising from interpersonal conflicts, disproportionate levels of participation and perceived unfairness of marks (Conway, Kember, Sivan & Wu, 1993; Sivan, Yan & Kember, 1995; Zhang, Johnston & Kilic, 2008). One way to manage these issues and learn from the processes of group work is through self and peer assessment (SPA) (Johnston & Miles, 2004; Knowd & Daruwalla, 2003; Zhang et al., 2008). SPA is a recognised pedagogic strategy that helps to identify individual contributions to group work and helps students to gain a better understanding of both the processes and outcomes of group work (Hughes & Large, 1993; Williams, 1992; Somervell, 1993).

A key challenge in managing the SPA process is the allocation and administration of marks and feedback, which lecturers use in their assessment and students can use in the development of transferable employability skills. Numerous studies have offered practical advice on best practice in managing group work and peer and self-assessment (cf., Conway et al., 1993; Goldfinch, 1994; Ireland, Jones & Ollin, 2003; Knowd & Daruwalla, 2003; Orsmond, Merry & Reiling, 2002 Sivan, 2000). Drawing on a Higher Education Academy – Hospitality, Leisure, Sport and Tourism Network funded project, this paper builds on and adds to this body of knowledge by examining the application of CASPAR (Computer Assisted Self and Peer Assessment Ratings) – a software tool that aids in the administration of SPA. The paper has three key aims: first, to offer critical commentary on CASPAR’s functionality, which can inform a) decisions about whether or not to adopt the system, and b) decisions about how it can be best deployed in learning and teaching; second, to provide critical discussion and guidance on SPA; and third, to offer discussion and guidance on the effective use of CASPAR within learning, teaching and assessment. In meeting these aims, the paper contributes to our understanding of how SPA can be improved, regardless of whether it is conducted using CASPAR or some other paper-based system, as well as how electronic resources can enhance learning, teaching and assessment strategies (cf., Batey, 2002; Biscomb, Davenport & Lane, 2008; Dale & Lane, 2004; Haven & Botterill, 2003; Lominé, 2002; McGugan, 2002).

WHAT IS CASPAR?

CASPAR is an internet based software tool developed to manage the assessment of group work more effectively (see CEMP, 2008 for details and demonstrations of the software and Lugosi, 2009 for user advice). CASPAR allows lecturers to set up group projects and students to participate in online SPA throughout the life of the project. SPA can take place once or on multiple occasions and alternative marking criteria can be set up for different assessment points to coincide with specific phases of group projects. It also allows students to provide qualitative feedback and quantitative marks, and it has a number of features that help to facilitate the group work process e.g. project journal and
live messaging service. Lecturers can monitor progress, moderate marks and collect feedback, which can be used within formative and summative assessment.

CASPAR V1 emerged as a result of the HEFCE funded Group Work Assessment in Media Production (GWAMP) project (2000-4), which identified the need for more effective methods of managing the SPA process (see CEMP, 2007). CASPAR V1 was a separate follow-on project to GWAMP, developed within and funded by the Centre for Excellence in Media Practice (CEMP) at Bournemouth University (CEMP, 2008). CASPAR V1 used “.ASP”, a Microsoft based software technology. Following a small-scale pilot test of CASPAR V1 (2005-6), a second version was developed using “PHP MySQL”, a free, open-source software technology. CASPAR V2 has required live testing with end users to identify and fix technical and functionality issues. This testing phase was accompanied by an evaluation of the online SPA process to establish good practice for its use in learning and teaching. This paper discusses some of the themes and data that emerged through this evaluation and offers critical commentary on the application of CASPAR.

THE STUDY

The study aims and objectives

The overall aim of the original study was to assist in the testing, evaluation and finalisation of CASPAR V2 prior to global distribution. Four key objectives were set: firstly, to evaluate the experiences of staff using CASPAR in their teaching and in the management of group work; secondly, to evaluate the experiences of students using CASPAR in group work and assessment; thirdly, to provide user feedback to support testing, validation and finalisation of CASPAR V2; and finally, to provide guidance on the effective use of CASPAR in group work and SPA.

Context

The evaluation part of the study took place in the School of Services Management, Bournemouth University. Lecturers and students from six subject areas (Hospitality, Leisure, Sport, Tourism, Retail and Events) participated in the study. CASPAR was used on eight units (four second year undergraduate, three final year undergraduate and one masters), which had in total 288 enrolled students.

Data collection methods

The primary method of formal data collection was questionnaires, distributed on paper and electronically via the Bristol Online Survey system. Students were given a questionnaire that focused on four areas: 1) personal information, including age, sex,
nationality, course, level and unit on which CASPAR was used, 2) the use and functionality of the system, 3) the norms of assessment, including the establishment of the marking criteria and 4) students’ appraisal of the SPA/feedback process. From the total target population (n=288) a 50.69% response rate was obtained (n=146).

Questionnaire responses were complemented by feedback from users: students communicated their concerns and emerging issues verbally and through email to the relevant tutors; feedback from lecturers piloting CASPAR was also gathered through informal conversations throughout the project and a semi-structured interview following their use of CASPAR. Responses from lecturers focused on three particular sets of issues: a) the context and processes of SPA, including the assessment aim and the establishment of the marking criteria, b) emerging technical and practical issues encountered during assessment, which included the setting up of CASPAR for assessment and processing information, and c) the overall experiences of using CASPAR. Information about technical and practical issues was relayed during the project to the developer, who made necessary changes and updates to the system.

Problem formulation and analysis

Initial feedback from tutors about their teaching and assessment highlighted that the establishment of the marking criteria and the SPA was conducted slightly differently on the second year Production and Service Operations Management (Ops Man) unit than on the other units. Within the Ops Man unit the marking criteria was established collaboratively with students and the SPA process was taught alongside concepts of empowerment (cf. Lashley, 2001). Moreover, students were encouraged to use SPA as a developmental process and encouraged to identify one area in which their peers were good and one area where they needed to improve. The aim was to encourage reflective learning and peer-led self development. Colleagues on other units also stressed the developmental aspect of SPA but consultation on the marking criteria and SPA introduced as part of the teaching of empowerment were not part of the learning and teaching strategies for the other units from which student responses were gathered. Frustratingly, students from one unit were given the opportunity to participate in the development of their assessment criteria, but none of those students participated in the survey. Nevertheless, it was assumed that the different approach to SPA was an influencing factor that shaped students’ perceptions of both the SPA process and its outcomes. This led to the identification of specific hypotheses, which informed the subsequent analysis. More specifically, it was assumed that students on the Ops Man unit felt more empowered and engaged in the development of the assessment process. Consequently, two hypotheses were proposed:

H1. Students on the Ops Man unit will display a more positive attitude toward the marking criteria briefing than students on other units.

H2. Students on the Ops Man unit will display a more positive attitude toward the marking criteria than students on other units.
Furthermore, because students on the Ops Man unit were actively encouraged to engage in a process of critical reflection and to use the SPA as a developmental process, three further hypotheses were proposed:

H3. Students on the Ops Man unit will be more truthful in their peer assessment than students studying on other units.

H4. Students on the Ops Man unit will find the process of writing qualitative feedback more helpful in learning what makes good group work than students studying on other units.

H5. Ops Man students will display a higher perceived ability to work in a group than those students studying on other units as a result of using CASPAR.

Following the production of descriptive statistics, the sample was split into two groups: one group, which used CASPAR on the Ops Man unit (n=65), and a second group, which included students from all the other units (n=81). The Mann-Whitney test was used to measure differences between the two groups, and where statistically significant differences were observed ($p < .05$), and the null hypothesis was rejected, the nature of the differences was examined and effect size ($r$) was calculated (see Field, 2005). Following Cohen (1992), Field (2005, p.32) argued that the effect size is a “standardised measure of magnitude of the observed effect”; $r = .10$ suggests low effect; $r = .30$, medium effect and $r = .50$, large effect. Highlighting statistical significance alongside effect size in this case helps to provide a more objective evaluation of the impact of the different teaching strategies on students’ perceptions of the SPA process.

Limitations

The use of CASPAR was determined by the assessment schedule and the use of peer assessment and group work in the Spring and Summer term of the 2007-8 academic year. The group assessed work on many units in the School was already completed and it was therefore not possible to trial CASPAR. Nevertheless, CASPAR was used on a range of units from five different subject areas.

A further challenge in the data gathering was the response rate, which was just over 50%. If the exercise was repeated, more pressure would be placed on students to complete the questionnaires. In future it may also be useful to design a questionnaire into the system’s functions, which users have to fill in to complete the assessment.

FINDINGS AND DISCUSSION

The system

The testing of the CASPAR system in the eight units helped to identify a number of areas that needed further revision. The majority of these were technical issues, for example, for a short period in one unit the feedback was not anonymised by the system, the system
was inaccessible at times, the graphs showing performance over time were not functional and the interface page for entering marks and comments occasionally slipped out of line. These were corrected during this evaluation project and CASPAR V2 is now available for general use (see CEMP, 2008).

Use and functionality

70.5% (n=103) of student respondents claimed they found it easy or very easy to access CASPAR, and 66.4% (n=97) claimed they found it easy or very easy to use the system. Only 9.6% (n=14) found it difficult to access CASPAR and 12.3% (n=18) found it difficult to use it. In the majority of those cases, the difficulties encountered were technical rather than functional. In other words, occasionally the system encountered technical difficulties and did not operate properly; but when it was operating the functions were easy to use. These technical problems have been resolved. No respondents had extreme difficulties either accessing or using the system. However, several students found to their detriment that once submitted they could not edit comments or marks.

Lecturers similarly noted that problems were largely technical rather than functional. However, there were three challenges faced by users. Firstly, lecturers required administrative support in gathering student information, which has to be uploaded to the system’s database using a spreadsheet document. To do this, staff have to be assigned an admin status and have to enter the data in a “.csv” file format (see CEMP, 2008 for details). The system can allocate students to groups randomly or alphabetically, or lecturers can manually select students. In principle the allocation of students by CASPAR saves time, although this is unhelpful when groups are self-selecting and when lecturers are faced with a large cohort of students. This was even more difficult when students were already assigned to seminar groups, and lecturers had to select students individually for inclusion into specific work groups. Several lecturers noted that they found this process of uploading data and allocating students to groups a time and labour intensive exercise. Lugosi (2009) offers some practical advice for managing this process, including uploading the seminar groups as separate .csv files and setting up projects within seminar groups, which may have three of four groups, rather than having a single project for an entire cohort of 100+ students and having 25 or more groups to allocate. Nevertheless, both methods involve a great deal of effort.

Secondly, once a project is started it is not possible to move students to different groups or change marking criteria or assessment dates. Lecturers therefore had to be very careful in setting dates as well as checking spelling and grammatical accuracy. Miscalculating the sequencing of activities, for example setting an assessment date before a key milestone in a group project rather than after, made SPA on that part of group work invalid. Finally, lecturers could not discard a particular assessment. This became a problem with the initial SPA when several students submitted incorrect marks/inappropriate or incomplete comments. The time-intensive nature of setting up CASPAR and the inflexibility of the system to accommodate changes made several of the lecturers reluctant to use CASPAR again. Some have decided to use CASPAR in the subsequent academic year, but others have opted for paper-based SPA.
Assessment procedures

As noted previously, with the exception of Ops Man unit, students did not participate in the setting up of the marking criteria. However, 39.7% (n=54) of respondents said that in the future they would like to participate actively in setting the marking criteria, 44.9% (n=61) remained neutral, while only 13.9% (n=19) of respondents said they would not like to be involved. There was no significant difference between the Ops Man and other students in their response (p =.206). The majority of respondents (62.9%, n=90) felt the marking criteria were explained adequately and most respondents (69.4%, n=100) agreed with the marking criteria. 8.3% (n=11) of respondents disagreed with the marking criteria and 16.1% (n=23) felt the criteria were not explained in enough detail prior to assessment.

It is interesting to note that students on the Ops Man unit differed slightly from those on other units in their agreement with the marking criteria ($U = 1946.50, p$ (2-tailed) =.004, Z = -2.86, n = 144, r = -.24) and in their attitudes towards explanations of the criteria ($U = 1820.50, p$ (2-tailed) =.002, $Z = -3.10, n = 143, r = -.26$). Ops Man students had a greater tendency to agree with the marking criteria and to be more positive about the explanations of the criteria, although it is important to highlight that the effect size is moderate to low (Field, 2005) and it is therefore important to remain cautious in rejecting the null hypotheses for H1 and H2.

This does, however, highlight the potential for using the process of setting up CASPAR, and peer assessment in general, in the development of key transferable employability skills highlighted by Yorke (2004). In the Ops Man unit, CASPAR had been piloted over two years rather than one. The Unit has a practical and a theoretical element. Students studying this Unit have to design operational plans for a commercial training restaurant, which includes menu planning, production planning, marketing, food production, service and data analysis, which is used to drive further improvement. In both years, students were asked to design their own assessment criteria for the development of the operational plan. The involvement of students in the design of the assessment had three purposes: firstly, as noted previously, this consultation exercise was conducted during the teaching of empowerment and thus illustrated to students the practicalities of different empowerment strategies (cf. Lashley, 2001). Secondly, it gave students the opportunity to engage in the entire assessment process. They could consequently develop a more nuanced understanding of the exact criteria used to assess their performance.

Thirdly, it can help students to think critically about what are assessable activities or traits and how those measures of assessment can be articulated and operationalised. Initial suggestions from students on the Ops Man unit often included very general categories of assessment, for example, “attitude”, “performance”, “contribution.” Through discussions in lectures and seminars these were refined and in some cases abandoned and replaced by more specific criteria. These included: 1. Contribution to group discussions (i.e. the amount, not the quality); 2. Reliability in carrying out allocated roles/tasks; 3. Quality of written work/work produced; 4. Acceptance of advice and criticism, including follow-up actions; 5. Punctuality; 6. Reliability in attendance; 7. Organisation and preparation.
for meetings (this includes knowledge of relevant operations management principles).

Involving students in the SPA process in this way can encourage the development of meta-cognition (i.e. learning how to learn), the development of essential employability skills needed for industry (Dochy, Segers & Sluijsmans, 1999; Macphearson, 1999), and consequently the emergence of reflective practitioners capable of critically evaluating the social and organisational worlds they inhabit (see Lashley, 1999; Tribe, 2002). Therefore, it also responds to current calls for students to develop a broad set of employability skills (People 1st, 2005; Yorke, 2004) and for embedding employability in the curriculum (Yorke & Knight, 2006).

Peer and self-assessment

The peer assessment in the different units was conducted anonymously and 86.1% (n=123) of respondents felt peer feedback and marking should be anonymous. Just over half (53.9%, n=77) claimed they did not hold back from writing/marking truthfully in assessing their peers because they were worried about receiving negative marks or feedback. Roughly the same number of people (51.1%, n=73) claimed they did not hold back from truthful peer assessment because of concerns that these would upset their colleagues. The results suggest that Ops Man students were more truthful and less inhibited than their colleagues in other units in their peer assessment. The fear of lower marks ($U = 1716.00, p$ (2-tailed) =.001, $Z = -3.46, n = 143, r = -.29$) or the fear of upsetting colleagues ($U = 2002.00, p$ (2-tailed) =.023, $Z = -2.27, n = 143, r = -.19$) had less influence on their marking and feedback; but, again, the effect size is moderate to low and it may be unwise to reject the null hypothesis for H3 despite the statistical significance. 60.3% (n=75) of all respondents agreed with the feedback they received from their colleagues, although only 45.3% (n=63) acted on the feedback.

Over half of those surveyed (65.5%, n=93) agreed with the statement that they ‘found the process of writing qualitative feedback helpful in learning what makes good group work.’ Only 6.3% (n=9) disagreed and 5.6% (n=8) disagreed strongly with this statement, while 21.8% (n=31) claimed they neither agreed nor disagreed. Interestingly, there was no statistically significant difference in the responses of the two groups and it was not possible to reject the null hypothesis for H4.

The results bring into focus the usefulness of stressing to students that they should provide feedback for all their colleagues, and that they should be encouraged to highlight positive aspects alongside areas where they can improve. Student responses also reemphasise the usefulness of engaging in SPA within a broader process of reflective learning and development. As Moon (1999, 2004a) has argued, reflection provides important opportunities for growth and specifically the development of key employability skills (Moon, 2004b). Emphasising to students the role of SPA in the development of these skills may also help those unsure or unaware of their learning to appreciate the usefulness of such exercises. Furthermore, although many students claimed they were truthful in their assessment, it is apparent that peer assessment may still be perceived by many to be divisive. Addressing these concerns during briefing sessions with students
can reemphasise the need to be mature and sensitive in providing collegiate feedback. Moreover, it is useful to stress that feedback should be seen as an opportunity (for future self development) rather than a threat (to current status).

The use of CASPAR and SPA within a broader process of reflective learning and skills development was reemphasised by students’ responses about whether the use of CASPAR had improved their ability to work in a group. 27.1% (n=38) felt that using CASPAR improved their ability, but for a further 37.9% (n=53) it was not clear whether it helped or not. 26.4% (n=37) disagreed with the notion that it helped, while 8.6% (n=5) disagreed strongly. Ops Man students demonstrated a greater tendency to agree with that it had improved their ability, although the effect size was once again moderate to low and it is therefore important to remain cautious in drawing conclusions and rejecting the null hypothesis for H5 ($U = 1800.00$, $p$ (2-tailed) = .007, $Z = -2.71$, $n = 140$, $r = -0.23$). One question raised by this result is whether students actually did not benefit from using CASPAR, or SPA, or whether they were unaware of their development. If students are simply unaware then making the potential learning outcomes for this exercise more overt is a key challenge for lecturers. If it does not help them to develop employability skills, then lecturers have to question whether it is simply used as a panoptical technique of surveillance and control (Foucault, 1991), through which students can assert power over colleagues. As Tan (2004) argued, self assessment, and by extension peer assessment, subverts students within regimes of normality and compliance. CASPAR and SPA in general may therefore be deployed purely instrumentally to maintain a social order. Alternatively, they may simply be used as pedagogic tools that provide further insights into the contribution of particular students to group work.

Whether they are used as a pedagogic tool for fairer allocation of marks, a vehicle for reflective learning or a form of surveillance, the majority of respondents (69.4%, n=93) felt that peer assessment should take place at multiple points throughout the group project, with 30.6% (n=41) claiming that peer assessment should take place at the end. Providing students with several assessment points helps to generate formative feedback throughout the life of the group work project. Moreover, students also have opportunities to address emerging issues as well as to improve their results.

CONCLUSION

This paper emerged from an evaluation of an electronic tool, CASPAR, although many of the points raised here about the processes of its use in the Ops Man unit are applicable to SPA generally. However, it is useful to think about SPA through CASPAR as a particular pedagogic strategy that presents its own set of challenges and opportunities. Setting up projects and allocating groups is time and labour intensive, and the inflexibility of the system may actually be the source of extra stress for lecturers. Seen this way, SPA through CASPAR is not necessarily more efficient than a paper-based system; it appears to displace rather than reduce work. However, once set up, CASPAR allows for multiple points of assessment and the use of different criteria at those assessment points; students receive instant feedback through qualitative and quantitative indicators, and they can also use various functions of the software, i.e. the project journal and the
instant messaging, to facilitate more effective group work. Collecting, collating and distributing marks and feedback through a manual method, for example a paper-based approach, would require much greater levels of lecturer input.

The effectiveness of SPA and CASPAR, like that of any other pedagogic tool, is determined by how it is used within learning, teaching and assessment strategies. It is interesting to note that the students studying on the Ops Man unit, and who consequently used SPA/CASPAR as part of a broader reflective learning approach, had a greater tendency to agree that the marking criteria was clearly explained and they tended to feel more positive about the criteria used in their assessment. Ops man students were also less likely to be inhibited in their peer marking and feedback. They also showed a greater tendency to react positively toward the statement that their ability to work in a group had improved through SPA using CASPAR. It is useful to stress that the effect sizes of the different group responses were medium to low, and therefore it is important to remain cautious in drawing conclusions. Nevertheless, if they are valid, they highlight the value of SPA through CASPAR within a multifaceted pedagogic strategy.

SPA using CASPAR can provide insights into individual performance and group dynamics. It can also shed light on individuals’ perceptions of their own capabilities, as well as the perceptions of their peers. For students they can provide a fairer and more transparent method for allocating marks for different group members. However, SPA and CASPAR can also be used within a more ambitious development process. As well as giving students a sense of empowerment, it can also be used to demonstrate the opportunities and challenges offered by different empowerment strategies. A critical awareness of this is undoubtedly an important quality for graduates. Moreover, both the processes and the outcomes of SPA using CASPAR (i.e. marks and qualitative feedback) can be used within a broader strategy of reflective learning through which students can develop key employability skills.

REFERENCES

Batey, J. (2002). Web page implementation and cultural change within a first year undergraduate module. *Journal of Hospitality, Leisure, Sport and Tourism Education, 1*(1), 51-60. DOI:10.3794/johlste.11.15


