Validating the assessment of individuals within undergraduate teams.

Authors:

Peter Willmot, Loughborough University, Loughborough, UK, Tel +44 1509 227555
p.willmot@lboro.ac.uk
Adam Crawford, Loughborough University, Loughborough, UK, Tel +44 1509 227198
a.r.crawford@lboro.ac.uk

Abstract *— Of all the problems associated with student learning in a team situation, the difficulties of precise assessment are supreme. Academics who feel comfortable setting examinations and individual coursework assignments are deterred from devising team assessments because they cannot guarantee that lazy students might benefit from the efforts of their teammates or that weaker team members might dilute the efforts of the more diligent individuals. We draw comparisons between undergraduate project teams and teams outside the university and discuss how accurately we can recreate the conditions of industry. We examine the argument for giving equal marks to all team players. We then set out to search for evidence of the wisdom or otherwise of applying a simple peer review to team marks and how fairly such a review can assess the individual team members.*

A web-based system of self and peer assessment, devised by the authors, has been in use at Loughborough for a number of years. We compare accumulated data with the opinions of 'fly on the wall' mentors: fourth year students attached to each project team in a supervisory role. The mentors know the capabilities of individual students particularly well and could offer an unbiased assessment. We also consider previous academic track record as an alternative means of verification.

The paper then goes on to consider if peer marking can inform us about more than just the performance of individuals. A low standard deviation in peer marks may be a sign of harmony in a team while a high standard deviation might indicate disharmony. We look for evidence of where an apparently harmonious team has raised their performance above the normal expectations its members. While results were mixed with no absolute relationships being established, the results indicate some interesting trends and point the way to a better understanding of this very complex subject.

Index Terms *— Peer review, peer moderated, self-assessment, team projects, web based,*

INTRODUCTION

Academics who feel comfortable setting examinations and individual coursework assignments are deterred from devising team assessments because they cannot guarantee that they can assess the individuals within the teams accurately. A lazy student might benefit from the efforts of his or her teammates or particularly diligent students may have their efforts diluted by weaker team members. Failure to address this very real problem may not only cause resentment amongst the students but is also considered quite unacceptable from the Quality Assurance point of view.

Studies by working groups of academics on behalf of a recent joint funded programme on project based learning based at Nottingham and Loughborough Universities distilled current practice in UK higher education [1]. Most acknowledge the problem and that many of the mechanisms in place to address it are less than ideal. Self and peer assessment has been the focus of much of the recent work in this field where marks or weightings, collected from the team members themselves, are used to modify a team mark allocated by the project supervisor through various mechanisms. This paper sets out to examine the wisdom or otherwise of applying a simple automated self/peer marking method known as Web-PA and to examine some of the pedagogical issues associated with it.

WELCOME TO THE REAL WORLD

Some would say that if we are to mimic the 'real world', the emphasis in engineering education should be firmly centred on the team. It might be argued that in industry just as on the sports field, the outcome is a team effort and all benefit equally from the team's success. When your favourite football team wins the championship, for example, even the substitutes receive a winner's medal because it is thought to be a team effort. Is it fair, therefore to conclude that teamwork at university should be integral in the curriculum and that in the case of team assignments, individuals must be prepared to entrust their future to the collective outcome? How strong is this argument for giving equal marks equal marks to all team players?

Being in a team has become an inescapable feature of modern professional life. We are all 'team players'. In the work context, the received wisdom is that teamwork delivers real organisational benefits; improving productivity, reducing absenteeism and enhancing employee satisfaction. The outcome of an effective team is expected to exceed the

sum of its component parts. In the wider world, teams have a clear, highly positive image, embracing the likes of sporting teams like Manchester United or the New York Yankees, or maybe in the field of music in a brass band or symphony orchestra. A place in such a team confers clear status: where you compete to get into the team and have to perform consistently well if you are to keep your place.

This paints a highly optimistic picture of teamwork, however, and although we may be able to point to some high achieving student teams that all pull together effectively, it is sure that some colleagues in academia will have had very different experiences working with other, less cohesive groups of students. When it comes to work-based project teams, the reality is that project managers don't have the luxury of choosing the members, they take who they can get. It is a lottery too, for the team members. So perhaps we didn't ought to concern ourselves over much about team selection in degree programmes provided we mix it up now and then. Some colleagues argue that we should take personality profiles like Myers Briggs type indicators or Belbin team roles into account, but we find that this is neither practical nor does it reflect what happens in industry. Our experience indicates that an element of self-selection tied in with an element of compulsory intermixing and team rotation is the best route to assembling satisfied teams to represent the 'real world' scenario.

Other differences between a sports team and a project team are obvious in the role and power of the manager and the importance of practice. A sports team is managed by someone external to the team; the club coach. A conductor manages an orchestra. These individuals are the arbiters of standards and their decisions are final. Project managers have to operate in a more complex managerial environment, which inevitably limits their powers and discretion. It is in this area, however that the student project team veers towards the sports team and consequently away from the real world of engineering industry. In recent years we have very successfully experimented with mentors (who are degree finalists) attached to each junior student team. It appears that the mentors are often more able to effectively communicate ideas successfully than the teaching supervisor who is in a more authoritarian position. A strong parallel can be seen here between the industrial project team and it's section leader reporting to a senior manager who is inevitably more detached from the day-to-day work.

A sports team is given time to practise, and an orchestra to rehearse; project teams are, somehow, expected to gel and perform from day one. Why then should we expect our students to hit the ground running at their first experience of teamwork? Our anecdotal experience shows that students get better at teamwork with practice and that, while teamwork instruction has its place, practice is more effective. This is particularly evident in the speed with which a newly formed team with experienced members goes through its evolutionary stages as described by Tuckman [2] and begins to get down to real work. It is clear that students benefit from tackling various types of teamwork throughout their degree programme.

Clearly, the different types of teams under discussion work in radically dissimilar environments, although it can be argued that it is the project team environment that is the most complex to analyse. Unlike the football analogy, it seems each project is unique and the conditions for team selection and motivation are often less than ideal. Academics should not feel isolated in this; it is just a matter of degree. As for drawing parallels in academic assessment of individual players; the highly successful sports team will command the greatest accolades and even the weakest member of that team will demand a share in that good fortune. Similarly all players in a relegated football club can expected to take a hit. On the other hand, the brightest stars within the cup-winning club can expect further enhanced salaries and advertising contracts while the stars in the relegated club may collect a lifeline transfer. It is likewise in industry; a successful project team is judged by that project's success or failure and the strong or weak team members can expect marginal adjustments. This sits well with the idea of overall team marking modified by incremental adjustments for individuals but the question for academics is how best to make those adjustments.

WEB BASED PEER REVIEW

An innovative web-based system of self and peer assessment has been in place at Loughborough across a number of modules and disciplines for several years. It provides a convenient mechanism through which individuals can be assigned a proportion of the supervisor mark based upon data entered in confidence by the members of the teams. The system, known as Web-PA was designed to handle large numbers of students in a flexible, tutor driven manner and was described at ICEE2004 [3]. The system is currently undergoing further development as a result of local user feedback with a view to extending its capabilities and making it more widely available for ongoing research into the subject through the newly formed Engineering Centre for Excellence in Teaching and Learning (UK) which is based at Loughborough. We shall shortly be seeking partners around the world to take part in this work. Very briefly, Web-PA allows teamwork supervisors to set up a number of criteria against which team members mark themselves (self) and the other team members (peers) at the end or at any other prescribed time in a project. Students enter data in confidence from any web terminal and the system calculates a variation factor for each team member based on the total score for an individual divided by the average scores for the whole team. The supervisor marks the team effort in the usual way and this mark (or a part of it) is multiplied by the factor for each individual. Where all team members score equally, the Web-PA factor is 1.0 and all members gain the nominal team mark.

INDEPENDENT VALIDATION

Clearly it is always going to be difficult to obtain independent validation of Web-PA given that the peer review process is invoked to inform the team mark because we lack knowledge of the precise contributions of each student. In the research published at ICEE2004, we showed that the Web-PA algorithm compares well with others that are used elsewhere and dissected the Web-PA output data to see how well self marking compared with the marks awarded by the other members of the team. In most teams we found considerable agreement about which members were the main contributors. However, it is not surprising that an individual's perceptions of their contribution, truthful or otherwise, is often higher than the group's (peer) perception with the biggest difference occurring in the weakest group members who probably enter false data to avoid failing the module. Nevertheless, the system was found to consistently and quite correctly punish weak members and reward stronger ones, even though the differences in marks appear less pronounced than the perceived wisdom would suggest they should have been. This does give the assessor a degree of comfort as it means students are seldom actually failed by the application of Web-PA. Furthermore, as we discussed earlier, this outcome does appear to accord with the real world scenario that even weak members are able to claim the rewards for being part of a successful team.

A new approach to validation considers the opinions of final year M.Eng (Master of Engineering) students for the first time this year. Second year (level 2) mechanical engineering undergraduate students routinely tackle a team design project at Loughborough. Students visit an industrial company from a small consortium where they are introduced to a real problem. A finalist works as 'mentor' to each team of three or four. Each consortium company sets the problems for four or five teams (16-20 students in all) and an academic supervisor takes charge of the activities of that company group and assessment of the project outcomes in consultation with a company tutor. In 2004/5 there were eight company groups operating with more than 30 teams. Projects run from mid October to early May on one afternoon per week. The novel mentoring experience forms part of a 'Leadership' module for the finalists and was described at ICEE2003 [4].

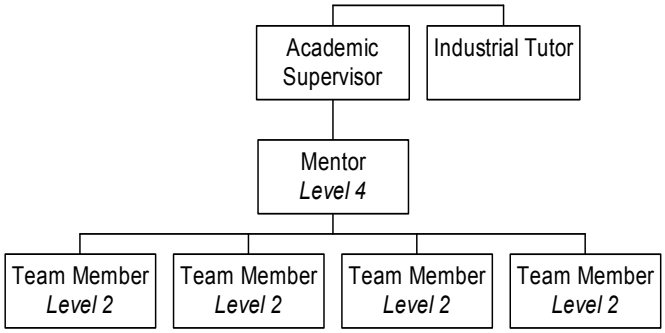


Figure 1, SUPERVISORY STRUCTURE

Mentors build strong working relationships with their teams through weekly meetings, some of which are observed by the supervisor but the age and maturity gap ensures that the relationship remains on a professional level. Mentors focus on the task and are not directly involved in team assessment. Mentors are marked on their performance as a leader, how they manage the project and deal with problems that might occur. They write a reflective critique at the end of the project and therefore get to know the individual students very well; their strengths and weaknesses. They are consequently in an excellent position to offer candid and unbiased opinions.

At the end of the project, this year, mentors were each issued with a paper version of the peer review completed by the second year students asking them to score their team under the same criteria used in Web-PA.. Mentors were advised that the information was being gathered for anonymous use into peer review research. Completing the form was optional and 22 out of 27 mentors responded. The process assumes that students can score marks by contributing to any or all of the published criteria and the peer assessment factor is calculated by summing the marks under each heading. The six criteria used were:-

- Ability to find and retrieve technical information
- Ability to generate Ideas and concepts
- Ability to methodically evaluate concepts
- Contribution to modelling and development of design solutions
- Attendance and diligence.
- Communication skill

Figure 2 (overleaf) is a plot of the mentor response against the weighting factor calculated by Web-PA for the design project in academic year 2004/5. The mentor scores having been faithfully manipulated in the same way as in the web tool, that is; mentor assessment = score for the individual divided by average score for all team members. Of the 59 students for which we have data, 85% show broad agreement between the Web-PA result and the opinion of the mentor (defined as within the parallel lines drawn at +or- 0.1 either side of the line of equality). Of the nine students who fell outside these error bars, six are located in the first and third quadrants (both values above unity or both below) which suggests that both response mechanisms agreed that the individual was a strong/weak team member but disagreed about the degree of strength or weakness. More worrying are the three (1.5% of the total) that appear in the second and fourth quadrants indicating considerable disagreement.

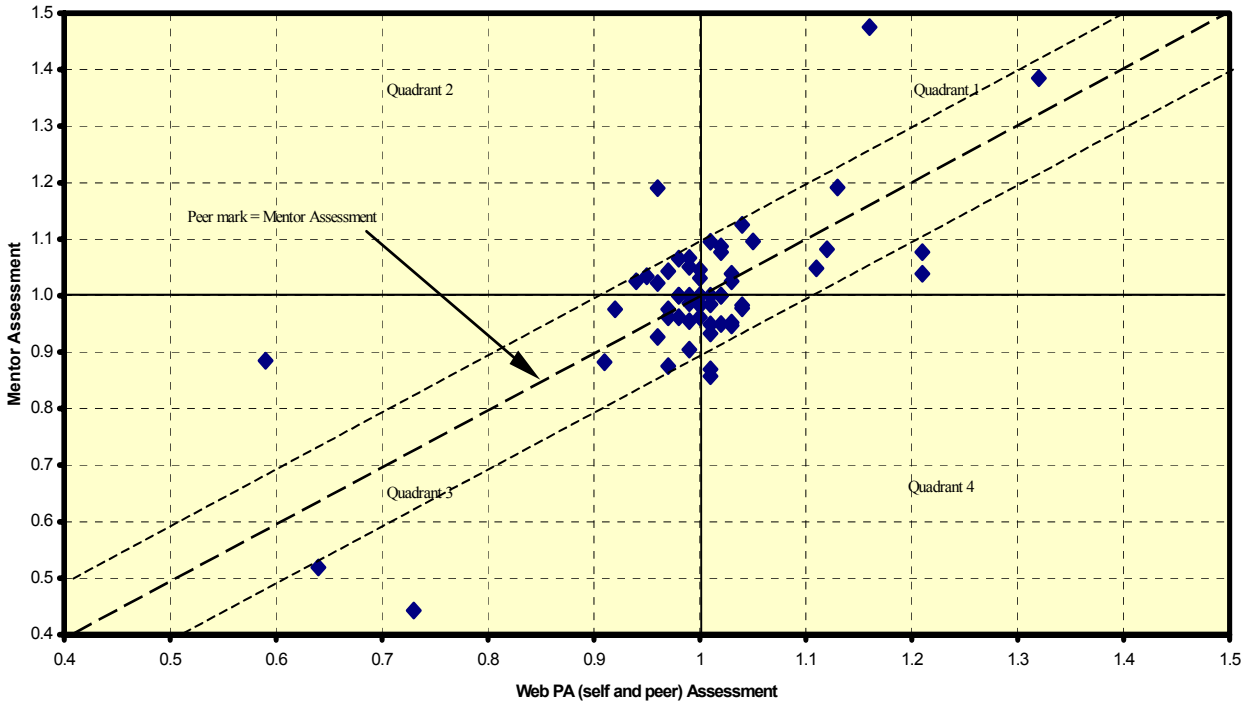
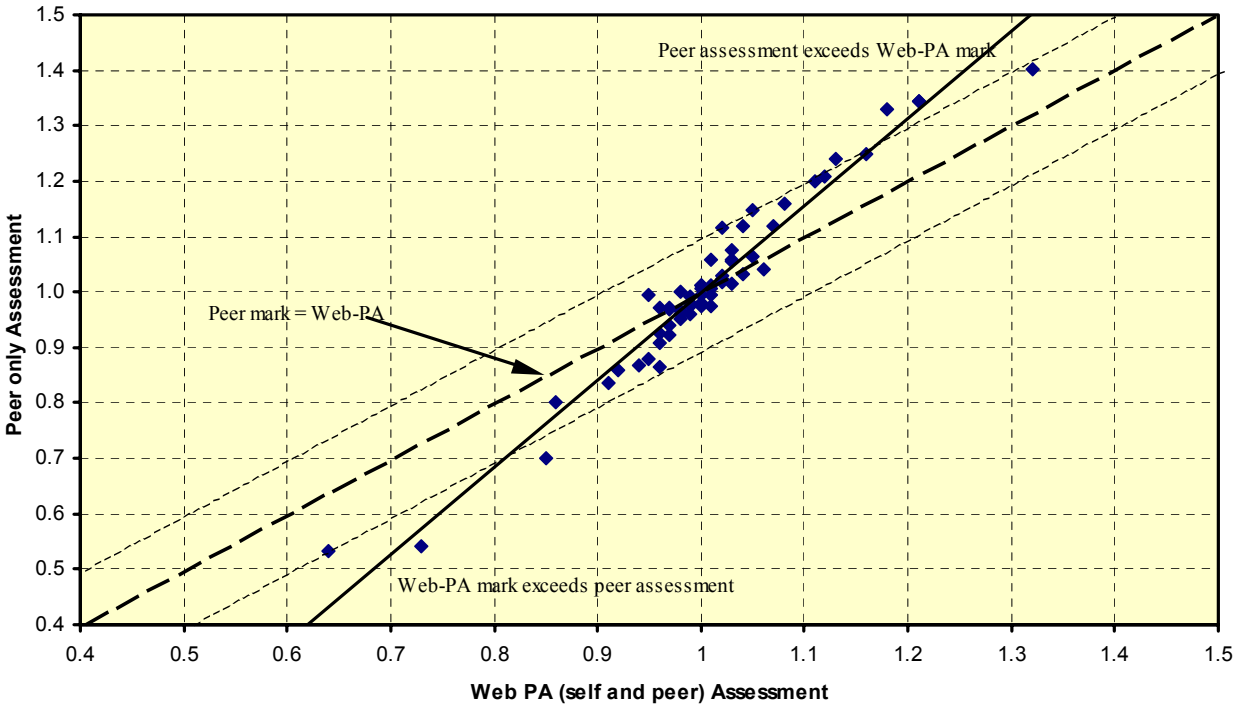


Figure 2, COMPARING WEB-PA OUTPUT WITH MENTOR ASSESSMENTS

It is interesting to consider the effect that members marking themselves has on the Web-PA scores: for the same year group (and including those teams for which no mentor assessment was available) the self-assessment was removed. Figure 3 shows this output – the opinions of the peer group only – plotted against the normal Web-PA output. A clear trend is discernable: the Web-PA system has the effect of moderating the extreme of marks. Strong students still receive added marks but less so than would be suggested by peer marking only and weak students are afforded a certain protection. From the course leader's viewpoint this offers some reassurance that students will not be overly disadvantaged.



FIGUR E 3, COMPARING WEB-PA WITH PEER ONLY MARKS

Closer examination of the data sets reveals the unsurprising fact that a large number of students award marks in higher proportion to themselves than their peers do. However it was interesting that plots of assessments including both self-marks and peer marks (the Web-PA scenario) actually provided a better correlation with mentor assessments than when either marks were compared in isolation with the mentor assessment. Although this study is limited in size, it does give some reassurance of the validity of using a review system that includes both self and peer-marking.

SEARCHING FOR EVIDENCE THAT TEAMWORK REALLY CAN DELIVER ACADEMIC SUCCESS

As Belbin[5] said, "Nobody's perfect, but a team can be" An effective team is not just a band of outstanding players; a combination of skills are needed but the key is how the team members work with each other. As previously stated, the basic objective of good teamwork should be to achieve synergy. The authors speculated that synergy might be equated with harmony and that a harmonious team might therefore be expected to outperform a team in conflict with a similar skills base.

Taking year marks from academic transcripts over two completed years as the benchmark for the ability of any given student, we compared these figures with a considerable amount of peer assessment data from the past projects of second year students. We searched for trends which might indicate that harmonious teams (teams with a small variation in peer marking) had exceeded the academic expectations of the team members. The results, however, were inconclusive. Overall year marks based on a large number of modules that include examinations and a variety of coursework assignments tend to be consistently lower than marks for the team project activity in any case. Most students like industrially based team projects and they tend to work particularly hard at them. Even allowing for this, it is not difficult to identify harmonious teams that had performed beyond reasonable expectations. On the other hand, we were equally able to identify many teams that had displayed disharmony but had still done better than one might expect. A few showed the exact opposite characteristic. We had to conclude that it is unreasonable to use overall academic performances as an absolute indicator of how well engineering students might tackle open-ended design projects. After all, experience has frequently shown up students who are not particularly strong in the examination room but who excel in a creative team situation. Close inspection of the data shows some of the largest variations in peer marks occurred where three members were approximately equally rated but the fourth was shown up to be weak – an indicator by our metrics of disharmony. It is perhaps not surprising, on reflection, that the team project mark was still quite good: the three would take over the task to all intents and purposes perhaps allocating minor assignments to the weak member and the team of three would continue to perform as a sub-unit.

Furthermore when Web-PA assessment was compared with an equivalent factor, calculated from the year marks, that relates the relative academic strength of each team member to the team average, the conclusion was that team project performance is often very different from performance the academic mainstream. An example of the scattered results is shown in figure 4

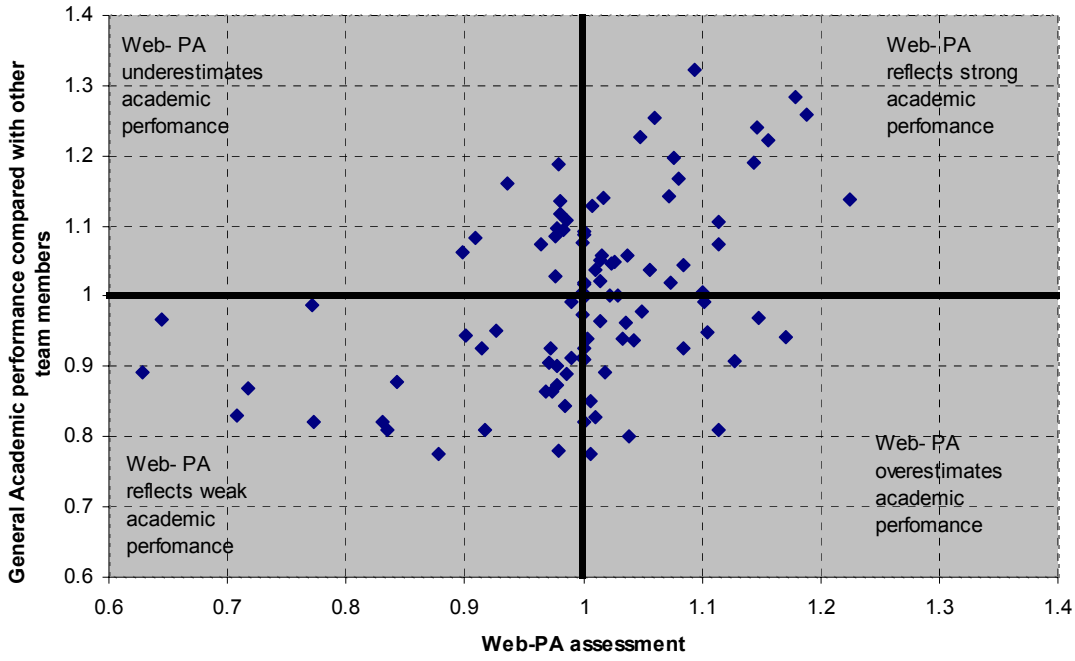


Figure 4, Web-PA RESULTS COMPARED WITH THE RELATIVE ACADEMIC ABILITY OF INDIVIDUALS

Although not, perhaps what we were looking for when we embarked upon this research, the evidence appears strong that first class students are not necessarily first class in respect of team projects and some students who perform quite poorly in examinations contribute well to team projects. This raises some interesting further questions.

The team projects that form the subject of this study were real world problems set exclusively by industrial companies. The industrial tutors were in touch with th teams throughout and were present at review meetings along the way. They also double marked the ritten reports along with university supervisors. It is logical, therefore to conclude that this project work is about as close as one can get in a university situation to the kind of task an engineer in industry might tackle. Traditional assessment methods do not appear to provide a yardstick by which we can measure a student's ability to tackle such a project so are they still the most appropriate to determine whether our students are adequately prepared for life in industry? Or are real world projects the more accurate determinant of who will perform well upon graduation?

CONCLUSIONS

- Self and peer assessments submitted privately by team members against stated criteria broadly agrees with the opinions of 'fly-on-the-wall mentors.
- On this sample only 1.5% of the results showed widely divergent opinions. This number could be explained by personality clashes or by mentors misjudging the abilities and efforts their students
- Peer only-marking tends to over exaggerate differences between the team members and a combined system of self and peer marking tends to moderate extreme marks and provides a comfort zone for the supervisor.
- The research has posed further questions about the validity of traditional examination methods as accurate indicators of the ability of university students to perform in industry.

REFERENCES

[1] Crawford, A,R, Tennant, J, Choo, B, S, Wilson A (Ed). "A Guide to Learning Engineering through Projects", *FDTL(3) report 43/99 PBLE: Project Based Learning in Engineering* edited works, Sept 2003. http://www.pble.ac.uk

[2] Tuckman, B, 'Developmental sequence in small groups' Forming Storming Norming and Performing model, 1965

[3] Willmot, P, Crawford, A, R, "Online peer assessed marking of team projects" *Proceedings of the International Conference on Engineering Education, ICEE 2004, Paper WA7/4* , Gainesville, Florida, USA, October 2004, pp 1-7

[4] Willmot P, "Introduction of Student Mentors into a Programme of Industry Linked Projects" , *Proceedings of International Conference on Engineering Education, ICEE 2003* , Area 7Paper 1367 , Valencia, Spain, July 2003, pp 1-5 .

[5] Belbin M, "Management Teams - Why they Succeed or Fail*"*, 2003 - 2nd edition, Butterworth Heinemann, Oxford