

**Leniency and halo bias in industry-based assessments of student competencies: a critical, sector-based analysis**

Dr Katharina Wolf

*School of Marketing, Curtin University, Perth, Western Australia*

GPO Box U1987, Perth, Western Australia, 6845. [K.wolf@curtin.edu.au](mailto:K.wolf@curtin.edu.au)

# **Leniency and halo bias in industry-based assessments of student competencies: a critical, sector-based analysis**

Industry placements are popular means to provide students with an opportunity to apply their skills, knowledge and experience in a 'real world' setting. Within this context, supervisor feedback allows educators to measure students' performance beyond academic objectives, by benchmarking it against industry expectations. However, industry assessments appear to be frequently clouded and overwhelmingly positive by nature, which questions the reliability and validity of supervisors' judgment of competencies.

Supervisor bias has been paid much attention within the context of clinical placements, as well as within the domain of social work. However, the concept has been largely ignored within business education, despite the increasing emphasis on - and deep integration of - work integrated learning in the business curriculum.

This paper sets out to address this gap by examining variances in mark distribution and apparent leniency in the context of a final year, compulsory placement unit, based on observations and data collected over nine semesters (n=546). The focus of this study is less on the reliability of existing measures, than on gaining an understanding of the reasons behind assessment bias and the pressures placed on industry assessors. The data indicates that different types of placement locations apply dissimilar standards when assessing student performance. The author identifies three statistically different placement types (small business, not-for-profit and professional), which influence the strength and risk of grader bias, hence ultimately the assessment outcome.

Keywords: assessment, internship, placement, leniency, industry, halo, business education, public relations

## **Introduction**

Higher Education institutions are under pressure to provide the industry with *job ready* graduates, who require minimum training and understand what is required of them in the so-called *real world*. Institutions who can demonstrate that their learning outcomes reflect the skills and attributes desired by the industry will in return be equipped with a

powerful marketing advantage, as their graduates are more likely to be employed (Rundle-Thiele, Bennett, & Dann, 2005).

However, beyond making employment decisions, to what extent should industry representatives be involved in the assessment of industry-readiness and the allocation of marks? Should their role be limited to that of a mentor, industry supervisor and advisory group member, or should they actively participate in providing students with quantitative, for credit feedback that will enable a realistic understanding of industry expectations? Students enrolled in the public relations (PR) internship unit, which provides the focus for this study, regularly argue that supervisor feedback should be “the only” component contributing to their placement mark, as it allows to measure industry readiness rather than the achievement of academic objectives. However, in comparison to internal marks, industry assessments appear to be frequently clouded and overwhelmingly positive by nature, which challenges the reliability and validity of supervisors’ judgment of competencies.

This phenomenon is investigated in this paper, by not only comparing students’ internal, academic marks against evaluations provided by their industry supervisors, but more importantly by comparing feedback across different types of placement locations.

### **The role of work integrated learning in public relations**

The focus of this paper is on the Public Relations (PR) Major taught at Curtin University’s Business School, Australia, as part of the Commerce Degree. The Major is offered across five different campuses throughout Southeast Asia and in Mauritius. The course is positioned as an industry-focused, practical degree, which prepares students for a smooth transition into an exciting, fast paced and ever-changing industry. As part of the curriculum, students complete a compulsory, final year placement, covering a minimum of 20 days in a communications-focused position.

Practica and compulsory work placements have long formed a crucial element of the higher education curriculum, particularly in health sciences and education studies. However, while work integrated learning is becoming increasingly popular across disciplines, Kerr and Proud (2005) discovered that *real life* experience and industry placements are particularly emphasised by the public relations industry, even more so than in other, related business and communication disciplines. Compared with traditional professions, such as law and medicine, public relations is still in its infancy. Consequently, there is increased global pressure to ensure the professionalisation of the industry, which ultimately accentuates the need for stronger and more successful partnerships between academia and industry. As a result, potential employers, industry advisory committees and placement supervisors have become crucial stakeholders in the educational process. Their active involvement in the course design is a promise frequently made to attract students, demonstrating the real-world validity of the program.

Work integrated learning, and industry placements in particular, are highly popular amongst students, who frequently refer to internships as the most significant and enjoyable part of their undergraduate program (Bates, 2004; Crebert, Bates, Bell, Patrick, & Cragolini, 2004; Wolf, 2008). These enable students to apply their skills, knowledge and experience in a 'real world' setting, by benchmarking their academic performance against industry expectations. A professional industry placement provides students with an opportunity to gain a *glimpse of the real world*, while still benefiting from the relative safety and support provided by their university (Wolf, 2008). It offers them an opportunity to apply their professional skills and knowledge in an authentic environment, as well as to re-adjust their expectations and industry perception before entering the graduate employment market (Wolf, 2010).

For the purpose of this unit, supervisor feedback is collated at the end of the internship based on a 20-item questionnaire, which requires industry representatives to rate students' performance with the help of a seven-point Likert scale (see appendix). These items broadly cover the university's graduate attributes, with a key focus on professional conduct, communication skills and subject knowledge. Although the supervisor appraisal is considered a crucial assessment component, it is limited to 20% of the overall unit mark. The major contributor to the final unit mark is a comprehensive report, in which students are expected to critically analyse their placement experience (70%). This is supported by their active participation in a reflective group blog (10%). Students frequently challenge the perceived low impact of the supervisor appraisal on their final mark, whilst academic staff has noted its typically clouded, inflated nature. A preliminary study (Wolf, 2012) suggested consistently high levels of leniency and inflated marks when comparing supervisor appraisals with internal marks across five campuses over four semesters. While internal results were comparable to the usual course average around the mid-60s, industry assessments were on average over 20% higher. Considering the exaggerated nature of supervisor allocated assessment marks in what is essentially positioned as a capstone unit, the question emerges of how realistic - and useful - this type of feedback is, particularly in terms of identifying levels of industry preparedness and potential skills gaps prior to graduation.

## **Literature review**

Scholars have challenged the widely established norm that student performance should exclusively be assessed by academic staff members (e.g. Rainsbury, Hodges, Sutherland, & Barrow, 1998). This common practice is arguably ironic in itself, considering that many universities' key objective is the 'industry readiness' of their graduates. However, despite challenging the academic as exclusive intellectual authority, studies have consistently demonstrated that supervisor assessments are not as reliable and valid as widely assumed (Gonsalvez & Freestone, 2007). In fact, Marks and Regehr (2005) discovered that clinical supervisors often choose not to fail students, despite judging their performance as unsatisfactory. This lack of validity has led to rater

bias in supervisory evaluations being referred to as a “pervasive problem” (Robiner, Saltzman, Hoberman, Semrud-Clikeman, & Schirvar, 1998, p. 50).

Supervisor bias and assessment leniency have been well documented in learning and teaching literature. However, despite the recognition that supervisors across disciplines are inclined to rate a disproportionate amount of students highly, the vast majority of empirical research has been conducted within the context of clinical placements (Bandiera & Lendrum, 2008; Dohrenbusch & Lipka, 2006; Dudek, Marks, & Regehr, 2005; Karasik, 2009), psychology (Gonsalvez & Freestone, 2007; Hedge & Kavanagh, 1988; Robiner & Hoberman, 1997), and social work (Bogo, Regehr, Roxanne, & Regehr, 2007; Kadushin & Harkness, 2002; Lazar & Mosek, 1993; Pease, 1988; Vinton & Wilke, 2011). Arguably, this research focus may not be a major surprise, considering the serious – potentially life threatening - consequences of allowing insufficiently trained graduates to practice in their chosen, health related field. However, leniency and bias are equally common in business contexts (Rainsbury, et al., 1998), but have to date remained largely under-investigated. This study sets out to address this gap.

### ***Bias in student assessment***

The two key phenomena widely credited as responsible for inflated supervisor assessment are leniency and halo bias. Leniency bias, signifies the reluctance to assign low or fail grades to a student. In effect, some supervisors “tend to use only the upper or positive portions of an evaluation scale” (Bandiera & Lendrum, 2008, p. 45). According to research, fieldwork instructors perceive the evaluation of students as one of the most challenging, stressful and worrisome aspects of placement supervision (Bogo, et al., 2007; Pease, 1988). Inflated or overly favourable assessments may be the result of a supervisor’s fear of ‘retribution’, or the challenging of results, which would require further justification and potentially lead to conflict (Bandiera & Lendrum, 2008; Dudek, et al., 2005). Bogo, et al. (2007) highlight that it may simply not be in the nature of some supervisors to provide what they perceive to be “bad news”, particularly if they personally subscribe to a ‘nurturing’ worldview, focused on empowerment. However, supervisors may also be simply unfamiliar with the grading requirements and reporting expectations (Dudek, et al., 2005), or inadvertently attempt to ‘please’ the university representative (Vinton & Wilke, 2011).

As a result, feedback and supervisor endorsements may arguably become meaningless. For example, in their 2007 study, Gonsalvez and Freestone determined that 83% of the students in the first and 77% in the second cluster received distinction and high distinction grades. The low percentage of credits did not only suggest a leniency bias, it furthermore questioned the value of supervisor appraisals in an environment where students have limited opportunities to set themselves apart (Gonsalvez & Freestone, 2007).

### ***The halo effect***

Whilst leniency refers to a rater's general disposition, *halo bias* is more individualistic and essentially personal. Halo bias clouds an instructor's judgement by metaphorically hanging a halo over the evaluations of a student, hence giving him or her a higher rating than deserved (Pease, 1988). Also referred to as 'halo error', recognition of the halo effect dates back to the early 1920s when Thorndike (1920) defined it as a "marked tendency to think of the person in general as rather good or rather inferior and to color the judgements of the [specific performance dimensions] by this general feeling" (p. 25). Based on Thorndike's definition, halo can be described as a within-rater phenomenon that results from the rater's subjective, favourable – or indeed unfavourable – impression of the ratee. Consequently, the halo effect, or 'global impression' (Dohrenbusch & Lipka, 2006), may influence the result to the student's advantage or disadvantage.

Bandiera and Lendrum (2008) emphasise the common nature of 'halo' by arguing that in most cases the assessment of a learner's performance in multiple roles will ultimately be determined by only two or three overriding perceptions of the learner. The final assessment may be swayed by an outstanding impressions made at the outset. Conversely, earlier mediocre performance may be entirely forgotten by the time the appraisal is due. Furthermore, some raters may prefer to award grades based on 'potential', focus on compatibility with their own personality or indeed rate a student positively because they did not challenge authority or create any problems (Pease, 1988). Studies discovered that supervisors may perceive their own students as particularly outstanding and gifted (Miller & Van Rybroek, 1988; Robiner & Hoberman, 1997). Balzer and Sulsky (1992) critique the halo concept for not being a reliable measure. However, in the context of this study the author is more interested in investigating the underlying reasons and potential implications for assessor leniency, rather than aiming to accurately measure the extent of it.

### ***Reasons behind leniency***

According to Lazar and Mosek (1993) the 'contamination' of feedback is largely due to the excessive role played by the relationship between supervisor and supervisee in the assessment process, which they argue results in a fundamentally flawed grading system. Ratings have been found to be consistently high when instructors evaluate in a face-to-face situation, compared to greater variance when the evaluation process takes place anonymously (Vinton & Wilke, 2011).

Numerous suggestions have been made in the related literature regarding the management of grader bias, ranging from increased awareness of common biases (Gonsalvez & Freestone, 2007) to the need for more support, training and guidance (Bogo, 2006; Dudek, et al., 2005; Robiner & Hoberman, 1997), as well as the active involvement of instructors in the development of performance criteria (Vinton & Wilke,

2011). Communication between practicum sites and programs have been identified as a key area for the improvement of the placement experience (Lewis, Hatcher, & Pate, 2005). Additionally, increased dialogue among supervisors may result in decreased bias (Robiner, et al., 1998).

Bogo, et al. (2007) emphasise that while tools for field evaluation have increasingly attempted to provide standardised, objective and impartial measures of performance, these may be in direct conflict with supervisors' personal and professional values. They may feel uneasy about the gatekeeper role they have been allocated, obliged to advocating normative standards of professional behaviour rather than respecting diversity and providing avenues for growth. On the other hand, supervisors may be indifferent about internship requirements, which have no real bearing on the primary mission of their organisation or project (Lewis, et al., 2005).

### **Research methodology**

A preliminary analysis of the placement data across unit locations over initially four semesters indicated consistently high levels of difference between internal unit assessments and supervisor appraisals across all locations in which the placement unit is delivered. Supervisor assessments were on average 21.46% higher than the internal mark (Wolf, 2012), which suggests some level of bias and leniency, in particular as similar graduate attributes are evaluated across all unit assessments. However, the basic comparison between internal and external (n=371) results did not allow for a deeper analysis of patterns and potential reasons for the bias. Hence, the need for a longitudinal in depth study was identified, which enables a critical examination of the types of placement organisations chosen, and how these may influence grader bias.

In order to deepen the understanding of how the choice of placement location may influence supervisor appraisals, this study consists of an analysis of variance of marks across different industry sectors, i.e. types of placement locations. This includes an in-depth analysis of all 20 assessment criteria items over a period of nine semesters, mapped against five core graduate attributes (Discipline knowledge, thinking skills, communication skills, lifelong learning skills and application of professional skills) and compared across six common placement location types: Communications consultancies (or agencies), government departments (or agencies), small businesses, not-for-profit organisations (or charities), corporations and event companies. These six categories typically capture placement choices and preferences by students enrolled in the PR Internship unit. While most of these categories are arguably transferrable to other business disciplines, i.e. capture the breath of placements undertaken across the faculty, *event*-related placements may be specifically relevant to public relations students. In the context of this unit, events represent a popular career choice or ambition and have therefore been included as a separate category in the original study design.

For access reasons, the analysis is limited to students enrolled at the Australian home campus. However, based on the insight gleaned during the preliminary study, these onshore results do not vary dramatically from those received at the university's overseas campuses. Furthermore, even though students are enrolled in Australia, their placements cover local, interstate and overseas-based work experience. Only cases with full data sets, i.e. complete intern appraisal forms and clearly identifiable placement locations, were included in this analysis, resulting in a total of 546 unique cases.

RQ: Do workplace assessors in different types of placement locations apply dissimilar standards when assessing student performance? If yes, what are some of the plausible explanations for these variations?

The data set has been analysed with the aid of IBM SPSS Statistics 19, utilising its general descriptive functions, plot means and one-way analysis of variance.

## **Findings & discussion**

The aim of this study is to critically examine potential reasons behind leniency in student assessments, with the intention of improving the understanding of factors impacting on grade reliability and validity within the context of a final year placement unit.

Results of the longitudinal study are broadly consistent with the findings of the preliminary study (Wolf, 2012), indicating an average disparity of 21% between internal and external marks. Arguably, the difference in grade distribution for the appraisal (external) and coursework (internal) component may have been influenced to some extent by better practical than coursework skills. However, a closer inspection of the results over time suggests that industry supervisors are lenient markers, unwilling to fail students. Out of 10,920 unique marks based on the 20 item criteria, only 42 results (0.38%) are below 4 on the 7-point Likert scale, i.e. less than 50%. More importantly, these results are limited to a total of eight intern assessments, i.e. 1.47% of the total number of cases under investigation. This substantiates Bandiera and Lendrum's (2008) findings that supervisors fail to utilise the full evaluation scale made available. Results for few students are consistent across all three assessment components (blog, report and industry appraisal). Instead, it is not uncommon for students to receive outstanding supervisor feedback, but to consequently fail internal assessment components.

In comparison with Gonsalvez and Freestone's (2007) study it appears as if supervisor bias may be even more emphasised in the context of this final year internship unit. Based on the data provided, 92.85% of students in this sample received a distinction or high distinction for the placement appraisal component of their mark. A staggering 80.92% of students received a high distinction. Converting the results to the grading system used at the University of Wollongong, where Gonsalvez & Freestone's



study was based<sup>1</sup>, results in only a minor reduction of "high achievers" (88.25% and 68.44% respectively). In fact, over the period of nine semesters, 36.2% of supervisors allocated an appraisal mark between 95 and 100%. Only one student failed the assessment, which was due to not meeting the required minimum duration of the internship.

It would be improbable to claim that the results are an indication of excellence, i.e. that all public relations students from this particular university perform outstandingly when they complete their final semester placement. 14.15% of students (n=77) were awarded a 100% mark by their supervisor, a result that is highly unusual within the (Australian) higher education context. This result may be an indication of the limited time spent on and consideration given to student assessment. It may also be the consequence of the halo effect, allowing a student's commitment, enthusiasm or performance in other areas of competency override the overall assessment. Based on conversations with supervisors and qualitative feedback provided on some appraisal forms, certain assumptions can be made. Over the semesters a number of supervisors contacted the unit coordinator to discuss concerns in some areas of performance, most commonly related to 'attitude' and 'enthusiasm'. However, at the end of the semester supervisors were visibly reluctant to allocate results towards the lower end of the scale. Instead, they appear to have *reasoned* why poor or mediocre performance in one area of competency could be justifiably inflated based on 'potential' and outstanding performance in other areas. Drawing on the literature discussed earlier, this may be because supervisors are keen to attract placement students in the future. A reputation for "tough feedback" may discourage prospective interns. Additionally, teaching staff may request further justification, which has not only time and resource implications, but forces supervisors to carefully substantiate their appraisal decision.

The industry or business sector in which a placement is completed may provide further insight into the reasons behind an apparently inflated supervisor assessment. For the purposes of this study and further comparative analysis, internship data has been grouped into six different types of locations, as detailed in the methodology section. The descriptive statistics obtained via SPSS 19 are revealing in themselves, indicating an average appraisal grade of 6.20 out of a maximum of seven, across all industry sectors and marking items, i.e. 88.57%, which is more than 20% higher than the usual internal mark average at this particular university.

---

<sup>1</sup> Whilst the University of Wollongong awards distinction for grades between 75-85, universities on the west coast of Australia award distinctions for results between 70 and 80. The same applies to high distinctions, commencing at 85 and 80 retrospectively

Placement type	N	Mean	Std. Deviation
Consultancy (Cons)	82	5.9254	.77036
Government (Gov)	91	5.9765	.80317
Small Business (SmB)	96	6.5156	.56555
Not-for-profit (NFP)	130	6.2788	.70565
Corporation (Corp)	95	5.9843	.69075
Events (Ev)	52	6.4359	.80011
Total	546	6.2032	.73996

Table 1. Average student score based on placement location

Table 1 highlights differences in the mean scores across the six identified placement types. Whilst small business, not-for-profit organisations and events companies commonly award marks above 6 out of a total of 7, i.e. above 85.7%, consultancies, government departments and corporations appear to be slightly more reluctant to award high distinctions.

#### *Evaluation of Graduate Attributes*

For the purpose of this study, the 20 items on the supervisor feedback questionnaire were re-coded into five of the university's Graduate attributes, namely Discipline Knowledge, Thinking Skills, Communication Skills, Lifelong Learning Skills, and Application of Professional Skills. The descriptive statistics indicate that supervisors consistently mark students lowest for their critical thinking skills (5.8878 total mean), with exception of government departments and not-for-profit organisations, both of which rank students' communication skills last (5.6538 and 6.0500 respectively). On the other hand, lifelong learning skills were consistently awarded the highest marks (6.4223 total mean), with exception of event-based placements, where supervisors rated students' professional skills marginally higher (6.5269 as opposed to 6.5000 average).

Based on the research question identified in the methodology section, the hypotheses for the main data analysis are as follows:

Ho: There is no difference in student appraisals between different types of placement locations

H1: There are variations in student assessments based on the type of placement location, - i.e. consultancy, government, not-for-profit, small business, corporation and events – which are not due to chance.

From a unit and course coordinator perspective, it is valuable to determine whether there is statistical evidence that the choice in placement location can influence student performance, or rather the level of leniency reflected in the supervisor appraisal mark. Analysis of variance is a procedure used to test statistically significant differences between three or more independent sample means, i.e. grades awarded across different types of placement locations. Hence, a one-way ANOVA, followed by Post Hoc multiple comparisons, using the LSD (Least Significant Difference) procedure, was identified as most suitable analysis technique and applied to the complete data set. For analysis purposes, the dependent variables are the above identified graduate attributes, all of which are continuous variables, based on a score between a minimum of 0 and a maximum of 7. The independent variable is the placement type (categorical), as identified above.

The multiple comparison of industry supervisor feedback across the marking criteria and industry sectors indicates no statistical difference between placements conducted in consultancies (Cons), government departments (Gov) and Corporations (Corp). Arguably, this is not surprising, as these three types of placement locations typically hire qualified public relations staff with a relevant higher degree qualification, i.e. in all three cases departments consist of dedicated communication professionals, which in turn impacts on their understanding of the industry, as well as their expectations of students, or graduates. Hence, these three location types may be merged under the “professional” label, reflecting their level of industry specific training, expectations and understanding. However, student feedback received in all three location types in the professional category varied statistically from marks allocated to students completing their placements within small businesses, not-for profit organisations or event companies, as illustrated in Table 2.

	SmB			NFP			Events		
	Mean Dif	Std. Error	Sig.	Mean Dif	Std. Error	Sig.	Mean Dif	Std. Error	Sig.
Cons	-.59021*	.08884	.000	-.35340*	.08424	.000	-.51050*	.08979	.000
Gov	-.53910*	.06727	.000	-.30229*	.06106	.000	-.45938*	.06852	.000
Corp	-.53124*	.05143	.000	-.29443*	.04299	.000	-.45152	.05305	.000

Table 2. Multiple comparison of industry feedback across placement types

In order to further confirm the statistically significant difference in grader leniency between types of placement locations, a Tukey HSD (honestly significant difference) test was conducted. This test is used to conduct pairwise comparisons amongst the individual mark means, with the familywise significance level set at .05. The results, as detailed in table 3, confirm the statistical differences between the professional locations' means (i.e. consultancy, government and corporation) and those for not-for-profits and small businesses / events. However, the differences in supervisor feedback between consultancy, government and corporation means were confirmed as insignificant, thereby further strengthening the argument to merge these three types of placement locations under the "professional" label.

The Tukey HSD indicates a further, third subset, with statistical difference in results between placement feedback received in not-for-profit organisations and that awarded by supervisors in small businesses. However, event-based placements failed to clearly fall into a single category. On reflection, this is no surprise. First, the event category contains the smallest number of cases in this dataset. More importantly, based on the number of employees, turnover and organisational structure, event organisations can be described as small businesses. A re-run of the Tukey HSD test, with merged small business and event placement categories confirms three statistically different placement types, based on average external results for internship students (table 3).

Placement type	N	Subset for alpha = 0.05		
		1	2	3
Cons	82	5.9254		
Gov	91	5.9765		
Corp	95	5.9843		
NFP	130		6.2788	
SmB(inc. events)	148			6.4776
Sig.		.910	1.000	1.000

Table 3. Industry types recoded (Tukey HSD subsets)

A detailed analysis on the basis of all five individual graduate attributes (see table 4) further confirms that placement supervisors in those locations newly grouped under the "professional sector" label rate their students consistently different to their peers in not-

for-profit organisations and small businesses (including event companies). This difference in student results across placement types suggests different levels of leniency in supervisor marks, with the highest level of leniency to be found in the (newly merged) small business category.

Skills / sector mean rating	Professional	NFP	Sm. Bus (inc. events)
Discipline Knowledge	5.7575	6.1476	6.4398
Critical thinking	5.6005	6.0354	6.1963
Effective communication	5.7757	6.0586	6.4194
Lifelong learning	6.2462	6.5078	6.6176
Professionalism	6.1494	6.33953	6.5381

Table 4. Graduate attribute mean ranking based on three statistically different placement types

The results therefore reject  $H_0$  at least partially, in demonstrating statistical differences between some of the originally identified sector types, which can be re-grouped into three statistically distinct types of placement locations, based on existing appraisal data. The next section briefly discusses potential underlying factors for marker leniency and motivating factors for supervisors based on the newly identified location types.

#### *Placement type 1: Small business (SmB)*

Of those students achieving top (100%) marks, 46.7% (35) were based in small business, including startup construction companies, boutique wineries, independent publishers and retail outlets. The physical size of the organisation undoubtedly shapes the intensity of the student-supervisor relationship, and possibly accentuates the halo effect on the official assessment. Furthermore, small businesses may have a demand for public relations, marketing and publicity; however, typically that expertise may not exist in-house, or represents a minor “add on” to another job description. Hence, PR placement students provide small businesses with an opportunity to benefit not only from a temporary growth in their workforce, but furthermore from access to current discipline-specific expertise and knowledge. Supervisors may be experts in their field or industry (e.g. sales, publishing, construction, general business skills), but rarely have an in depth understanding of what public relations entails and what they can expect from final year communications students. Hence, many of the items on the end of placement appraisal form, e.g. reference to the student’s understanding of public

relations principles and suitability for employment in the field, are challenging for the supervisor to answer and judge.

The key focus for supervisors in small businesses may therefore be less on providing a quality learning experience for the student, than on the benefits to the business. This is arguably reflected in the time that is spent on the intern appraisal form. There are numerous examples where supervisors raised concerns about students' attitude or understanding of their actual business, either in person with the unit coordinator, or on the actual appraisal form. However, despite those earlier comments, students were nevertheless awarded consistently outstanding quantitative feedback. This suggests limited interest in the academic assessment component and arguably highlights the impact of the intimate working relationship between supervisor and student.

#### *Placement type 2: Not-for-profit (NFP) organisations*

34.7% of *top achievers* (26) were based in not-for-profit organisations, all of which were small to medium-sized (i.e. no national head offices). Although the result mean is over three per cent lower than that for students based in small businesses, an average mark of 89.7% in the supervisor appraisal nevertheless suggests a high level of leniency (see table 1). Again, the close student-supervisor working relationship may impact on the final appraisal, as well as the fact that many not-for-profit organisations heavily rely on volunteers and only have limited discipline-specific expertise amongst their permanent workforce. Consequently, *gratefulness* may lead to an increased halo impact.

Nevertheless, a close analysis indicates a more sophisticated and considerate approach to the student appraisal. Overall, results are inflated, in particular in comparison to internal marks; however, NFP representatives were more conservative than their small business peers when judging placement students' subject-specific abilities. This is likely due to the fact that not-for-profits may not necessarily have a dedicated communications department, but usually employ at least one staff member who is exclusively responsible for marketing and communication tasks and who has undergone at least some level of discipline-specific training. From a unit coordinator perspective, NFP organisations provide a fantastic learning ground for placement students, as they expose interns to a wide range of projects and usually offer more responsibility and autonomy than their corporate counterparts. They also represent a popular career choice. Many PR graduates have moved into communication roles in the not-for-profit sector and are offering internships opportunities to current students. This affinity with the course, and first-hand intern experience, may however further influence the level of leniency observed in the appraisal.

Most importantly, not-for-profit organisations attract people with certain personality traits, who may arguably be more inclined to provide positive reinforcement and opportunities for growth. Based on Bogo's (2007) findings, they may feel uncomfortable in the allocated gatekeeper role. Like their small business colleagues,

not-for-profit supervisors may also be relatively unconcerned about the marking criteria, as the appraisal and skills assessed have limited bearing on their business objectives (Lewis, et al., 2005). From an educator's perspective, the major concern in both contexts is the danger that the organisation's service agenda may not only override the assessment, but ultimately students' educational needs.

### *Placement type 3: Professional sector*

Results of this study indicate that students can expect similar results, or possibly similar levels of rater leniency, for placements in communication consultancies, corporations and government departments. Despite representing three distinct career paths for many graduates, all three locations are characterised by dedicated teams of communication professionals. These teams may vary in size, however, each is typically managed by a senior individual, who has undergone discipline-specific training and understands the demands of the sector. This knowledge and experience shape expectations of placement students and result in the desire to provide them with a realistic understanding of what will be expected when they enter the industry as communications graduates. Some supervisors in the professional sector reportedly use the industry placement as a recruitment tool, aiming to identify potential future employees. However, many senior industry professionals and former graduates from this PR program perceive it as their duty to "give back to the industry", by providing emerging communicators with valuable, practical learning opportunities.

Despite professional training and industry specific experience, rater leniency nevertheless appears to represent a major challenge. For example, five of those students receiving a 100% appraisal marks were based within the government sector, including two in politicians' campaign offices (3-5 largely part time staff), which once again emphasises the impact of the intimate nature of some student-supervisor relationships. A further six students received outstanding (100%) feedback from their corporate supervisor. However, of these, two were based within the University, working as part of small project teams (3-5 staff) and one student was already employed by the organisation. Furthermore, a number of supervisors were PR graduates and had expressed their affinity, gratitude and connection to the degree. Although these factors may positively influence the quality of supervision and project insight provided to students, they may in turn impact on the final mark, as the supervisor recalls his or her personal experience (Pease, 1988).

### *Final thoughts*

Arguably, affinity and relatable personal experience may result in high quality mentoring, guidance and ultimately a memorable, insightful work experience opportunity. However, if supervisor appraisals - as indicated in this study - are inflated,

the question arises how valuable this assessment component ultimately is. Furthermore, lenient supervisor judgment is likely to lead to a graduate's inflated perception of his or her competence, counteracting the focus on industry readiness.

Although this study focuses on data collected as part of a final year placement unit within the public relations major, findings arguably provide insight into the wider business education context. For example accounting, information systems or management students may act in a more generic role if they complete their placement in a not-for-profit organisation or small business, compared to the discipline-specific focus that can be expected within most corporate or consulting contexts.

## **Conclusion**

This study suggests that the choice of placement location may influence student performance based on supervisor assessments. Findings suggest consistently high levels of grader leniency in the external appraisal of business students' placement performance, which arguably provide graduates with a false sense of industry readiness. Three statistically different types of placement locations have been identified, i.e. small businesses, not-for-profit organisations, and what has been labelled the professional sector. Grader leniency appears to be present across all three, but to varying levels, determined by the sector's internal needs and priorities, as well as supervisors' personality traits, discipline specific knowledge and training.

Supervisor feedback is intended to allow educators and most importantly students to measure their performance against industry expectations. However, this study indicates that this unit's current, quantitative approach to supervisor feedback is fundamentally flawed, as the *one size fits all* assessment criteria fails to take business contexts and objectives, as well as supervisors' personality, skillsets and priorities into account. Whilst emphasising the importance of work integrated learning, industry closeness and job readiness, internship assessment criteria typically match academic expectations and procedures. Ironically, these are frequently removed from the realities of working life in which performance is seldom assessed based on seven-point Likert scales and 20 non-negotiable, equally weighted, cross-industry performance measures. Hence, the quantitative feedback currently sought fails to objectively reflect the quality of the learning experience, as well as level of students' industry readiness.

Further first-hand insight is required to confirm supervisor reasoning across placement types, as well as to differentiate between leniency and halo effect, although qualitative, in particular informal feedback, indicates that both perform a major role in student assessment: While supervisors in the professional sector may be driven by a commitment to the industry, recruitment intentions and an interest in educating the next generation of industry professionals, not-for-profit organisations and especially small businesses may perceive interns as a valuable extension to the existing workforce, its knowledge-base and abilities. Hence, gratefulness for support provided may result in



overall lenient results and explain the elevated level of high distinction appraisals. Furthermore, the academically-driven, graduate-attribute-focused, standardised marking criteria assume discipline-specific knowledge and the ability to judge students' technical skills, as well as their suitability for the profession. Depending on the size and type of the placement organisation, not all elements of the marking criteria may be covered. This ultimately challenges individual supervisor's knowledge and skillsets, as well as their personal reputation. Hence, they may feel inclined to let overall impressions override areas where a student may have performed less satisfactorily (halo effect), or where they otherwise would be forced to make assumptions.

## References

- Balzer, W. K., & Sulsky, L. M. (1992). Halo and performance appraisal research: A critical examination. *Journal of Applied Psychology, 77*(6), 975.
- Bandiera, G., & Lendrum, D. (2008). Daily encounter cards facilitate competency-based feedback while leniency bias persists. *CJEM, 10*(1), 44.
- Bates, M. (2004). From knowledge to action and back again: Building a bridge. *Asia-Pacific Journal of Cooperative Education, 5*(1), 7-14.
- Bogo, M. (2006). Field instruction in social work. *The Clinical Supervisor, 24*(1-2), 163-193.
- Bogo, M., Regehr, C., Roxanne, P., & Regehr, G. (2007). When Values Collide. *The Clinical Supervisor, 26*(1-2), 99-117.
- Crebert, G., Bates, M., Bell, B., Patrick, C. J., & Cragolini, V. (2004). Developing generic skills at university, during work placement and in employment: graduates perceptions. *Higher Education Research & Development, 23*(2), 147-165.
- Dohrenbusch, P. D. D. R., & Lipka, S. (2006). Assessing and predicting supervisors' evaluations of psychotherapists—an empirical study. *Counselling Psychology Quarterly, 19*(4), 395-414.
- Dudek, N. L., Marks, M. B., & Regehr, G. (2005). Failure to fail: the perspectives of clinical supervisors. *Academic Medicine, 80*(10), 84.
- Gonsalvez, D. C. J., & Freestone, J. (2007). Field supervisors' assessments of trainee performance: Are they reliable and valid? *Australian Psychologist, 42*(1), 23-32.
- Hedge, J. W., & Kavanagh, M. J. (1988). Improving the accuracy of performance evaluations: Comparison of three methods of performance appraiser training. *Journal of Applied Psychology, 73*(1), 68.
- Kadushin, A., & Harkness, D. (2002). *Supervision in social work*: Columbia Univ Pr.

- Karasik, R. J. (2009). Learning From Internships in Gerontology and Geriatrics: Assessment and Program Evaluation. *Gerontology & Geriatrics Education*, 30(4), 300-316.
- Kerr, G. F., & Proud, B. (2005, 4-7 December). *Hiring Graduates: Perspectives From Advertising And Public Relations Employers* Paper presented at the ANZMAC 2005, Perth, Western Australia (Proceedings on CD ROM).
- Lazar, A., & Mosek, A. (1993). The influence of the field instructor-student relationship on evaluation of students' practice. *The Clinical Supervisor*, 11(1), 111-120.
- Lewis, B. L., Hatcher, R. L., & Pate, W. E. (2005). The Practicum Experience: A Survey of Practicum Site Coordinators. *Professional Psychology: Research and Practice*, 36(3), 291.
- Miller, R. K., & Van Rybroek, G. J. (1988). Internship letters of recommendation: Where are the other 90%? *Professional Psychology: Research and Practice*, 19(1), 115.
- Pease, B. B. (1988). The ABCs of social work student evaluation. *Journal of Teaching in Social Work*, 2(2), 35-50.
- Rainsbury, E., Hodges, D., Sutherland, J., & Barrow, M. (1998). Academic, employer and student collaborative assessment in a work-based cooperative education course. *Assessment & Evaluation in Higher Education*, 23(3), 313-324.
- Robiner, W. N., & Hoberman, H. M. (1997). Psychology supervisors' training, experiences, supervisory evaluation and self-rated competence. *The Clinical Supervisor*, 16(1), 117-144.
- Robiner, W. N., Saltzman, S. R., Hoberman, H. M., Semrud-Clikeman, M., & Schirvar, J. A. (1998). Psychology supervisors' bias in evaluations and letters of recommendation. *The Clinical Supervisor*, 16(2), 49-72.
- Rundle-Thiele, S., Bennett, R., & Dann, S. (2005). The Successful Preparation and Development of Future Marketing Professionals: A Recommended Methodological Framework. *Journal for Advancement of Marketing Education*, 7, 27-35.
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4(1), 25-29.
- Vinton, L., & Wilke, D. J. (2011). Leniency Bias in Evaluating Clinical Social Work Student Interns. *Clinical Social Work Journal*, 39(3), 288-295.
- Wolf, K. (2008). *A glimpse of the real world: an investigation into the perceived effectiveness of compulsory public relations internships from an industry and student perspective*. Paper presented at the World Association for Cooperative Education - Australian Collaborative Education Network Asia Pacific 2008 Conference, Sydney, Australia.

Wolf, K. (2010). "I wouldn't have gone out there on my own"- a critical investigation of the benefits and pitfalls associated with compulsory industry placements." *Asia Pacific Public Relations Journal*, 11, 63-78.

Wolf, K. (2012). "*Brooke has been an Absolute Angel*" - Leniency and the Halo Effect in Industry Based Assessment of Student Competencies. Paper presented at the International Conference on Cooperative and Work-integrated Education, Istanbul, Turkey.

**Appendix: Assessment criteria (quantitative section only)**

Please indicate the extent of your agreement with the following statements about the intern on a scale of 1 to 7 by circling the appropriate number. 7 means strongly agree and 1 strongly disagree. If you feel neutral or unsure about the statement, please circle number 4.

	Intern's Name: _____	Strongly Disagree					Strongly Agree	
		1	2	3	4	5	6	7
1.	The intern is enthusiastic about his/her work							
2.	The intern is hard working							
3.	The intern is punctual arriving for work and appointments							
4.	The intern is reliable							
5.	The intern's dress and personal grooming are appropriate for my organisation							
6.	The intern is well organised							
7.	The intern meets deadlines							
8.	The intern displays an appropriate level of initiative							
9.	The intern is adaptable							
10.	The intern is willing to learn							
11.	The intern learns quickly							
12.	The intern works well in a team							
13.	The intern is creative							
14.	The intern has excellent oral communication skills							
15.	The intern has excellent written communication skills							
16.	The intern has an excellent grasp of public relations principles							
17.	The intern's level of general knowledge is appropriate for employment in the PR field							
18.	The intern is able to think strategically							
19.	The intern has an understanding of ethical issues in public relations							
20.	The intern is well suited to a career in public relations							